RÉGRESSION: BASES DÉFORMÉES ET SÉLECTION DE MODÈLES PAR PÉNALISATION ET MÉTHODE DE LEPSKI.

GAËLLE CHAGNY[(*)]

Octobre 2011

Ce document est essentiellement un complément à l'article *Penalization versus Goldenshluger-Lepski strategies in regression estimation with warped bases.* Il contient les démonstrations détaillées des principaux résultats. Il ne contient par contre pas ce qui concerne les motivations et justifications des méthodes employées, ainsi que les comparaisons avec d'autres stratégies d'estimation. On se réfèrera pour cela à l'article cité. Enfin, ce document contient aussi la justification théorique du remplacement de la quantité inconnue dans la pénalité par un estimateur (cf. section 6.9).

Table des matières

[(*)]: MAP5 UMR CNRS 8145, Université Paris Descartes, France
email: gaelle.chagny@parisdescartes.fr.

Première partie 1. Introduction

1. Objectif

Le but est d'estimer f dans le modèle

$$Y = f(X) + \varepsilon.$$

avec

- X le design aléatoire, réel, de loi donnée par sa fonction de répartition G, supposée dérivable, de dérivée g supposée strictement positive.
- ε, variable aléatoire réelle, centrée, de carré intégrable, de variance notée σ^2,
- f la fonction inconnue de $\mathbb{R}$ dans $\mathbb{R}$. On supposera que $f \in L^2((a;b), \mathcal{B}_{(a;b)}, \lambda_{(a;b)})$ pour un certain intervalle non vide $(a;b) \subset \mathbb{R}$, où $\mathcal{B}_{(a;b)}$ et $\lambda_{(a;b)}$ désignent respectivement la tribu borélienne et la mesure de Lebesgue sur $(a;b)$.

Pour simplifier, on supposera que X est à valeurs dans $(a;b)$, presque sûrement, ce qui revient à dire que le support de g est inclus dans $(a;b)$. Le principe de la méthode utilisée est de commencer par estimer la fonction

$$h = f \circ G^{-1},$$

par sélection de modèles (construction d'une famille d'estimateurs par projection, et sélection d'un estimateur convenable dans cette collection, par méthode de Goldenshluger-Lepski, ou par pénalisation) puis de définir un estimateur pour la fonction cible f par

$$\hat{f} = \hat{h} \circ G \text{ ou } \hat{h} \circ \hat{G},$$

où $\hat{h}$ désigne l'estimateur de h, et $\hat{G}$ désigne la contrepartie empirique pour la fonction de répartition G dans le cas où celle-ci est supposée inconnue.

2. Inégalité de concentration utilisée

Proposition 1. *(Version intégrée de l'inégalité de Talagrand)*
Soient $n \in \mathbb{N}\backslash\{0\}$, $\mathcal{F}$ une classe au plus dénombrable de fonctions mesurables à valeurs réelles et $(X_i)_{i\in\{1,\dots,n\}}$ une famille de variables aléatoires réelles indépendantes. On définit, quel que soit $f \in \mathcal{F}$,

$$\nu_n(f) := \frac{1}{n}\sum_{i=1}^{n}\left(f(X_i) - \mathbb{E}\left[f(X_i)\right]\right).$$

On suppose qu'il existe trois constantes strictement positives M_1, H et v telles que :

$$\sup_{f\in\mathcal{F}}\|f\|_\infty \leq M_1,\ \mathbb{E}\left[\sup_{f\in\mathcal{F}}|\nu_n(f)|\right] \leq H,\ \textit{et}\ \sup_{f\in\mathcal{F}}\frac{1}{n}\sum_{i=1}^{n}\mathit{Var}(f(X_i)) \leq v.$$

Alors, quel que soit $\varepsilon > 0$,

$$(1)\quad \mathbb{E}\left[\left(\sup_{f\in\mathcal{F}}(\nu_n(f))^2 - 2(1+2\varepsilon)H^2\right)_+\right] \leq \frac{4}{K_1}\left\{\frac{v}{n}\exp\left(-K_1\varepsilon\frac{nH^2}{v}\right) + \frac{49M_1^2}{K_1C^2(\varepsilon)n^2}\exp\left(-\frac{\sqrt{2}K_1C(\varepsilon)\sqrt{\varepsilon}}{7}\frac{nH}{M_1}\right)\right\},$$

où $C(\varepsilon) = (\sqrt{1+\varepsilon}-1)\wedge 1$ et $K_1 = 1/6$.

Remarque : on peut utiliser cette inégalité pour certaines classes $\mathcal{F}$ de fonctions non dénombrables, par exemple si $\mathcal{F}$ est une partie dénombrable dense d'un espace métrique.

3. Démonstration : utilisation de l'inégalité de Talagrand, version Klein et Rio

3.1. **Première étape : Application de l'inégalité de Talagrand version Klein et Rio (2005).** L'objectif est d'obtenir un majorant pour

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}}|\nu_n(f)| \geq H + y\right),\ \text{où } y > 0.$$

On va utiliser pour ceci la dernière inégalité du théorème suivant, du à Klein et Rio [21].

Théorème 1. *Soient $\mathcal{S}$ un ensemble au plus dénombrable de fonctions mesurables définies sur $\mathcal{X}$ espace métrique polonais et à valeurs dans $[-1;1]^n$, $(\xi_i)_{i\in\{1,\dots,n\}}$ des variables aléatoires indépendantes identiquement distribuées à valeurs dans $\mathcal{X}$. On suppose que pour tout vecteur $s = (s^1,\dots,s^n) \in \mathcal{S}$, pour tout $i \in \{1,\dots,n\}$, $\mathbb{E}[s^i(\xi_i)] = 0$. On définit :*

$$Z = \sup_{s\in\mathcal{S} s=(s^1,\dots,s^n)}\sum_{i=1}^{n}s^i(\xi_i),\ \textit{et}\ V_n = \sup_{s\in\mathcal{S}}\ \mathit{Var}\left(\sum_{i=1}^{n}s^i(\xi_i)\right).$$

Alors, pour tout $\lambda > 0$,

$$\log\mathbb{E}\left[\exp(\lambda Z)\right] \leq \lambda\mathbb{E}(Z) + \frac{\lambda}{2}\left(2\mathbb{E}[Z] + V_n\right)\left\{\exp\left(\frac{e^{2\lambda}-1}{2}\right) - 1\right\}.$$

En particulier, en notant $\underline{v} = 2\mathbb{E}[Z] + V_n$*, on obtient, pour tout* $x > 0$*,*

$$\begin{aligned}\mathbb{P}\left(Z \geq \mathbb{E}(Z) + x\right) &\leq \exp\left(-\frac{x}{4}\log\left(1 + 2\log\left(1 + \frac{x}{\underline{v}}\right)\right)\right), \\ &\leq \exp\left(-\frac{x^2}{2\underline{v} + 3x}\right).\end{aligned}$$

Remarque : La deuxième assertion provient de la première en utilisant la méthode de Chernoff.

On applique ceci de la façon suivante : on prend $\xi_i = X_i$; pour $f \in \mathcal{F}$, on définit

$$s_f^i : x \mapsto s_f^i(x) = \frac{f}{2M_1}(x) - \mathbb{E}\left[\frac{f}{2M_1}(X_i)\right].$$

et on prend $\mathcal{S} = \left\{s_f = (s_f^1, \ldots, s_f^n),\ f \in \mathcal{F}\right\}$.
On a bien, pour $f \in \mathcal{F}$ et $i \in \{1, \ldots, n\}$, $\mathbb{E}[s^i(X_i)] = 0$, s_f^i à valeurs dans $[-1; 1]$ (par définition de M_1) et Z se réécrit $Z = \sup_{f\in\mathcal{F}}(n/2M_1)\nu_n(f)$.
Le théorème s'applique donc, et l'on a, quel que soit $x > 0$,

$$\mathbb{P}\left(\frac{n}{2M_1}\sup_{f\in\mathcal{F}}\nu_n(f) \geq \frac{n}{2M_1}\mathbb{E}\left[\sup_{f\in\mathcal{F}}\nu_n(f)\right] + x\right) \leq \exp\left(-\frac{x^2}{2\underline{v} + 3x}\right),$$

c'est à dire, pour tout $y > 0$, en prenant $x = ny/2M_1$, et en utilisant la définition de H,

$$\begin{aligned}\mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(f) \geq H + y\right) &\leq \mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(f) \geq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\nu_n(f)\right] + y\right), \\ &\leq \exp\left(-\frac{n^2y^2}{8M_1^2\underline{v} + 6M_1ny}\right).\end{aligned}$$

On majore ensuite $\underline{v} = 2\mathbb{E}[Z] + V_n$: d'une part, $\mathbb{E}(Z) \leq (nH)/(2M_1)$, d'autre part,

$$\begin{aligned}V_n &= \sup_{f\in\mathcal{F}} \operatorname{Var}\left(\sum_{i=1}^n \frac{f(X_i)}{2M_1}\right), \\ &= \frac{n}{4M_1^2}\sup_{f\in\mathcal{F}}\frac{1}{n}\operatorname{Var}\left(\sum_{i=1}^n f(X_i)\right), \\ &\leq \frac{nv}{4M_1^2}, \text{ par définition de } v.\end{aligned}$$

On a donc,

$$\underline{v} \leq \frac{nH}{M_1} + \frac{vn}{4M_1^2}.$$

Il vient ainsi,

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(f) \geq H + y\right) \leq \exp\left(-\frac{ny^2}{2(4M_1H + v) + 6M_1y}\right).$$

On utilise ensuite

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{f\in\mathcal{F}}|\nu_n(f)|\geq H+y\right)\\
&\leq \mathbb{P}\left(\left\{\sup_{f\in\mathcal{F}}\nu_n(f)\geq H+y\right\}\cap\left\{\sup_{f\in\mathcal{F}}|\nu_n(f)|=\sup_{f\in\mathcal{F}}\nu_n(f)\right\}\right)\\
&\quad+\mathbb{P}\left(\left\{\sup_{f\in\mathcal{F}}\nu_n(f)\geq H+y\right\}\cap\left\{\sup_{f\in\mathcal{F}}|\nu_n(f)|=\sup_{f\in\mathcal{F}}-\nu_n(f)\right\}\right),\\
&\leq \mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(f)\geq H+y\right)+\mathbb{P}\left(\sup_{f\in\mathcal{F}}-\nu_n(f)\geq H+y\right),\\
&= \mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(f)\geq H+y\right)+\mathbb{P}\left(\sup_{f\in\mathcal{F}}\nu_n(-f)\geq H+y\right),\\
&\leq 2\times\exp\left(-\frac{ny^2}{2(4M_1H+v)+6M_1y}\right),
\end{aligned}
$$

car le raisonnement fait ci dessus pour contrôler les déviations de $\sup_{f\in\mathcal{F}}\nu_n(f)$ s'applique de la même manière pour contrôler celles de $\sup_{f\in\mathcal{F}}\nu_n(-f)$.

3.2. **Deuxième étape : Modification type Birgé-Massart.** L'objectif est d'obtenir un majorant pour

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}}|\nu_n(f)|\geq\lambda+(\eta+1)H\right),\text{ où }\lambda>0,\ \eta>0.$$

On applique l'inégalité obtenue à la première étape avec $y=\lambda+\eta H$, et l'on va majorer le membre de gauche de l'inégalité, c'est à dire minorer la quantité suivante, en suivant la démarche employée par Birgé et Massart dans le corollaire 2 p.354 de [5].

$$
\begin{aligned}
\frac{y^2}{2(v+4M_1H)+6M_1y} &= \frac{\lambda^2+\eta^2H^2+2n\lambda H}{2v+8HM_1+6M_1\lambda+6M_1\eta H},\\
&\geq \frac{\lambda^2+2n\lambda H}{2v+6M_1\lambda+M_1H(8+6\eta)},\\
&= \frac{a+b}{c+d+e},\text{ avec des notations évidentes.}
\end{aligned}
$$

Or quels que soient $a,b,c,d,e>0$,

$$\frac{a+b}{c+d+e}\geq\frac{1}{3}\left(\frac{a}{c}\wedge\frac{a}{d}\wedge\frac{b}{e}\right).$$

En effet,

$$\begin{aligned}\frac{a+b}{c+d+e} &\geq \frac{a+b}{3(c\vee d\vee e)},\\ &= \frac{1}{3}\left(\frac{a+b}{c}\wedge\frac{a+b}{d}\wedge\frac{a+b}{e}\right),\\ &\geq \frac{1}{3}\left(\frac{a}{c}\wedge\frac{a}{d}\wedge\frac{b}{e}\right).\end{aligned}$$

Appliquer ici, ceci donne

$$\frac{y^2}{2(v+4M_1H)+6M_1y}\geq\frac{1}{6}\left[\frac{\lambda^2}{v}\wedge\frac{2\lambda}{M_1}\left(\frac{1}{6}\wedge\frac{\eta}{4+3\eta}\right)\right].$$

On utilise ensuite

$$\frac{1}{6}\wedge\frac{\eta}{4+3\eta}\geq\frac{\eta\wedge 1}{7};$$

en effet, on a d'une part $1/6\geq 1/7\geq(\eta\wedge 1)/7$, et d'autre part,

$$\frac{\eta}{4+3\eta}-\frac{\eta\wedge 1}{7}=\begin{cases}\dfrac{3(1-\eta)}{7(4+3\eta)}\geq 0 \text{ pour } \eta\leq 1,\\[2ex] \dfrac{4(\eta-1)}{7(4+3\eta)}\geq 0 \text{ pour } \eta\geq 1.\end{cases}$$

Ainsi, il vient

$$\frac{y^2}{2(v+4M_1H)+6M_1y}\geq\frac{1}{6}\left[\frac{\lambda^2}{v}\wedge\frac{2(\eta\wedge 1)}{7}\frac{\lambda}{M_1}\right],$$

et par suite,

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}}|\nu_n(f)|\geq\lambda+(\eta+1)H\right)\leq 2\exp\left(-\frac{n}{6}\left\{\frac{\lambda^2}{v}\wedge\frac{2(\eta\wedge 1)}{7}\frac{\lambda}{M_1}\right\}\right).$$

3.3. **Troisième étape : Intégration.** L'objectif est maintenant de parvenir à l'inégalité annoncée.

On veut majorer l'espérance de la variable positive :

$$X:=\left(\sup_{f\in\mathcal{F}}(\nu_n(f))^2-2(1+2\varepsilon)H^2\right)_+.$$

On suit ici la preuve du lemme 6.1 de [9] et on commence par écrire :

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty \mathbb{P}(X \geq t)dt, \\
&= \int_0^\infty \mathbb{P}\left(\sup_{f\in\mathcal{F}} (\nu_n(f))^2 \geq 2(1+2\varepsilon)H^2 + t\right) dt, \\
&= \int_0^\infty \mathbb{P}\left(\sup_{f\in\mathcal{F}} |\nu_n(f)| \geq \sqrt{2(1+\varepsilon)H^2 + 2(\varepsilon H^2 + t/2)}\right) dt, \\
&\leq \int_0^\infty \mathbb{P}\left(\sup_{f\in\mathcal{F}} |\nu_n(f)| \geq \sqrt{1+\varepsilon}H + \sqrt{\varepsilon H^2 + t/2}\right) dt, \\
&\quad \text{(en utilisant } \sqrt{2a+2b} \geq \sqrt{a} + \sqrt{b}).
\end{aligned}$$

A ce niveau, on peut utiliser l'inégalité de la deuxième étape avec $\eta = \sqrt{1+\varepsilon} - 1$ et $\lambda = \sqrt{\varepsilon H^2 + t/2}$. On en déduit,

$$\begin{aligned}
\mathbb{E}[X] &\leq \int_0^\infty 2\exp\left(-K_1 n \left\{\frac{\varepsilon H^2 + t/2}{v} \wedge \frac{2(\eta\wedge 1)}{7}\frac{\sqrt{\varepsilon H^2 + t/2}}{M_1}\right\}\right) dt, \\
&= 2\int_0^{t_0} \exp\left(-K_1 n\left\{\frac{\varepsilon H^2 + t/2}{v}\right\}\right) dt \\
&\quad +2\int_{t_0}^\infty \exp\left(\left\{-K_1 n \frac{2(\eta\wedge 1)}{7}\frac{\sqrt{\varepsilon H^2 + t/2}}{M_1}\right\}\right) dt,
\end{aligned}$$

où t_0 est fonction de tous les paramètres intervenants ici, et où l'on choisit la convention $\int_0^{t_0} = 0$ si $t_0 \leq 0$. Comme les intégrandes sont positives, on majore de toute façon chacune des deux intégrales par les intégrales de 0 à $+\infty$, et l'on utilise dans l'intégrande de la deuxième :

$$\sqrt{\varepsilon H^2 + t/2} \geq \frac{\sqrt{\varepsilon}H + \sqrt{t/2}}{\sqrt{2}}$$

pour en déduire :

$$\begin{aligned}
\mathbb{E}[X] &\leq 2\int_0^\infty \exp\left(-\frac{K_1 n}{v}(\varepsilon H^2 + t/2)\right) dt \\
&\quad +2\int_0^\infty \exp\left(-\frac{2K_1 n(\eta\wedge 1)}{7M_1\sqrt{2}}(\sqrt{\varepsilon}H + \sqrt{t/2})\right) dt, \\
&= 2\exp\left(-\frac{K_1 n}{v}\varepsilon H^2\right)\int_0^\infty \exp\left(-\frac{K_1 n}{2v}t\right) dt \\
&\quad +2\exp\left(-\frac{\sqrt{2}K_1 n(\eta\wedge 1)}{7M_1}(\sqrt{\varepsilon}H)\right)\int_0^\infty \exp\left(-\frac{K_1 n(\eta\wedge 1)}{7M_1}\sqrt{t}\right) dt;
\end{aligned}$$

et il reste à calculer les deux intégrales restantes : on a , pour toute constante $C > 0$:

$$\int_0^\infty \exp(-Ct)dt = \frac{1}{C} \text{ et } \int_0^\infty \exp(-C\sqrt{t})dt = \frac{2}{C^2}.$$

Ainsi,

$$\begin{aligned}\mathbb{E}[X] &\leq 2\left\{\exp\left(-\frac{K_1 n}{v}\varepsilon H^2\right)\frac{2v}{K_1 n} + \exp\left(-\frac{\sqrt{2}K_1 n(\eta\wedge 1)}{7M_1}(\sqrt{\varepsilon}H)\right)\frac{2\times 49 M_1^2}{K_1^2 n^2(\eta\wedge 1)^2}\right\},\\ &= \frac{4}{K_1}\left\{\frac{v}{n}e^{-K_1\varepsilon\frac{nH^2}{v}} + \frac{49}{K_1(\eta\wedge 1)^2}\frac{M_1^2}{n^2}e^{-\frac{\sqrt{2}K_1(\eta\wedge 1)\sqrt{\varepsilon}}{7}\frac{nH}{M_1}}\right\},\end{aligned}$$

ce qui est l'inégalité cherchée.

□

Deuxième partie 2. Estimation quand la loi du design est connue

4. Rappels

4.1. **Observations.** Les observations sont constituées des couples i.i.d. $(X_i, Y_i)_{i=1,\dots,n}$, avec $Y_i = f(X_i)+\varepsilon_i$; les variables ε_i étant i.i.d. centrées de variance σ^2, et les variables X_i ayant pour densité g, et fonction de répartition G, et indépendantes des couples précédents. L'objectif est d'estimer la fonction de régression f, en commençant par l'estimation de $h = f\circ G^{-1}$, en supposant que G est connue. L'hypothèse principale est l'appartenance de la fonction h à l'espace de Hilbert $L^2([0;1])$. Ceci va permettre de construire des estimateurs par projection sur certains sous-espaces de $L^2([0;1])$ pour cette fonction.

4.2. **Modèles.** Les modèles sont des sous espaces vectoriels de $L^2([0;1])$. On note la collection $(S_m)_{m\in\mathcal{M}_n}$, où le cardinal de $\mathcal{M}_n$ est fini, de cardinal dépendant du nombre d'observations n. Les hypothèses sont les suivantes :

[$\mathcal{M}_1$] Chaque modèle est de dimension finie D_m telle que $1 \leq D_m \leq n$.

[$\mathcal{M}_2$] Les modèles sont emboîtés, c'est-à-dire, pour tout $(m_1, m_2)\in\mathcal{M}_n^2$, tel que $D_{m_1}\leq D_{m_2}$, $S_{m_1}\subset S_{m_2}$. On note $(\varphi_j)_{j\in\{1,\dots,D_m\}}$ une base orthonormée de S_m $(m\in\mathcal{M}_n)$.

[$\mathcal{M}_3$] Il existe une constante positive ϕ_0 telle que pour tout indice $m\in\mathcal{M}_n$ et toute fonction $t\in S_m$, $\|t\|_\infty \leq \phi_0\sqrt{D_m}\|t\|$. Ceci équivaut à la propriété suivante : $\|\sum_{j=1}^{D_m}\varphi_j^2\|_\infty \leq \phi_0^2 D_m$.

On se contentera des modèles classiques : trigonométriques, polynômes par morceaux avec subdivision dyadique, ondelettes à support compact.

4.3. **Espaces de régularité.** On considèrera les espaces de Besov $\mathcal{B}^\alpha_{2,\infty}$, pour $\alpha > 0$.
Ceux-ci sont définis de la manière suivante : pour r un entier positif et $v > 0$ et une fonction à valeurs réelles t, on pose

$$\Delta_v^r t(x) = \sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}t(x+kv),$$

pour x choisi tel que $x + kv \in [0;1]$, $k\in\{0,\dots r\}$. Ensuite, pour $u > 0$, le module de régularité est donné par $\omega_r(t,u)_2 = \sup_{0<v\leq u}\|\Delta_v^r t\|$. Ainsi, on dit que la fonction t appartient à l'espace de Besov $\mathcal{B}^\alpha_{2,\infty}$ si t appartient à $L^2([0;1])$ et si, pour $r = [\alpha]+1$, $|t|_{\mathcal{B}^\alpha_{2,\infty}} = \sup_{u>0} u^{-\alpha}\omega_r(t,u)_2 < \infty$. L'introduction de ces espaces est justifiée par leurs bonnes propriétés d'approximation : on utilisera le lemme 12 de l'article de Barron, Birgé, et Massart *Risks bounds for model selection via penalization.*

5. Estimation non adaptative

On fixe un indice $m \in \mathcal{M}_n$, et on estime la projection orthogonale de la fonction h sur le modèle S_m.

5.1. Construction de l'estimateur.

On définit le contraste suivant :

$$\forall t \in L^2([0;1]),\ \gamma_n^G(t) := \|t\|^2 - \frac{2}{n}\sum_{i=1}^{n} Y_i t \circ G(X_i). \tag{2}$$

Justifions ce choix :

$$\forall t \in L^2([0;1]),\ \mathbb{E}\left[\gamma_n^G(t)\right] = \|t\|^2 - 2\mathbb{E}\left[Y_1 t \circ G(X_1)\right] = \|t\|^2 - 2\langle f \circ G^{-1}, t\rangle.$$

Ainsi, $\gamma_n^G(t)$ est l'équivalent empirique de $\|t - f \circ G^{-1}\|^2 - \|f \circ G^{-1}\|^2$ qui est minimal quant $t = f \circ G^{-1}$. Donc minimiser le contraste γ_n^G fournit naturellement un estimateur empirique de $h = f \circ G^{-1}$. On définit donc

$$\hat{h}_m^G = \arg\min_{t \in S_m} \gamma_n^G(t), \quad \hat{f}_m^{G,G} = \hat{h}_m^G \circ G. \tag{3}$$

Il en résulte un unique estimateur dont l'expression est la suivante :

$$\hat{h}_m^G = \sum_{j=1}^{D_m} \hat{a}_j^G \varphi_j,\ ,\ \text{with}\ \forall j \in \{1, \ldots, D_m\},\ \hat{a}_j^G = \frac{1}{n}\sum_{i=1}^{n} Y_i \varphi_j(G(X_i)). \tag{4}$$

5.2. Etude du risque. Comme $\hat{h}_m^G$ estime la projection h_m de h sur le sous espace vectoriel S_m, il est naturel de décomposer le risque en utilisant $f_m^G := h_m \circ G$: on décompose la perte de l'estimateur,

$$\|\hat{f}_m^{G,G} - f\|_g^2 = \left\|f - f_m^G\right\|_g^2 + \|f_m^G - \hat{f}_m^{G,G}\|_g^2,$$

avant de prendre l'espérance,

$$\mathbb{E}\left[\|\hat{f}_m^{G,G} - f\|_g^2\right] = \left\|f - f_m^G\right\|_g^2 + \mathbb{E}\left[\|f_m^G - \hat{f}_m^{G,G}\|_g^2\right].$$

On majore déjà le terme de variance :

Proposition 2. *Avec les notations ci-dessus,*

$$\mathbb{E}\left[\|f_m^G - \hat{f}_m^{G,G}\|_g^2\right] \le \phi_0^2 \mathbb{E}[Y_1^2]\frac{D_m}{n}.$$

Preuve
On notera dans la suite $a_j = \langle f \circ G^{-1}, \varphi_j \rangle$, pour $j = 1, \ldots, D_m$.

$$
\begin{aligned}
\|f_m^G - \hat{f}_m^{G,G}\|_g^2 &= \left\| \sum_{j=1}^{D_m} (a_j - \hat{a}_j^G) \varphi_j \circ G \right\|_g^2, \\
&= \left\| \sum_{j=1}^{D_m} (a_j - \hat{a}_j^G) \varphi_j \right\|^2, \\
&= \sum_{j=1}^{D_m} (a_j - \hat{a}_j^G)^2 \text{ puisque les } (\varphi_j)_j \text{ sont orthonormées.}
\end{aligned}
$$

Or,

$$
\begin{aligned}
\mathbb{E}\left[(a_j - \hat{a}_j^G)^2\right] &= \mathrm{Var}(\hat{a}_j^G), \\
&= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i \varphi_j \circ G(X_i)\right), \\
&= \frac{1}{n}\mathrm{Var}\left(Y_1 \varphi_j \circ G(X_1)\right) \text{ (caractère i.i.d des observations)}, \\
&\leq \frac{1}{n}\mathbb{E}\left[Y_1^2 (\varphi_j \circ G(X_1))^2\right],
\end{aligned}
$$

ce qui donne le résultat annoncé, en sommant sur j, et en appliquant l'hypothèse de connexion de normes.

□

Pour l'étude du biais, on suppose maintenant que l'espace S_m de dimension D_m est l'un des modèles suivant : espaces de polynômes trigonométriques, espaces fondés sur les polynômes par morceaux de degré inférieurs ou égaux à r, ou espaces fondés sur des ondelettes à supports compacts.

Proposition 3. *Supposons que $h = f \circ G^{-1}$ soit un élément de l'espace de Besov $\mathcal{B}_{2,\infty}^{\alpha}$, avec $\alpha > 0$ fixé ($\alpha \leq r+1$ dans le cas où S_m est un espace de polynômes par morceaux de degrés inférieurs ou égaux à r). Alors, il existe une constante C_α, ne dépendant que de α telle que*

$$\left\| f - f_m^G \right\|_g^2 \leq C_\alpha D_m^{-2\alpha}.$$

Preuve
On note déjà que $\left\| f - f_m^G \right\|_g = \|h - h_m\|$, où l'on rappelle que h_m est la projection de h sur S_m. Par conséquent, le terme à majorer est aussi $\inf_{t \in S_m} \|h - t\|^2$. Le résultat provient donc directement d'un résultat d'approximation, le lemme 12 de l'article de Barron, Birgé et Massart, *Risks bounds for model selection via penalization.*

□

De ces deux propositions, on déduit immédiatement le corollaire suivant :

Corollaire 4. *Sous les hypothèses de la proposition précédente, la majoration suivante du risque de l'estimateur* $\hat{f}_m^{G,G}$ *a lieu :*

$$\mathbb{E}\left[\|\hat{f}_m^{G,G} - f\|_g^2\right] \leq C_\alpha D_m^{-2\alpha} + \phi_0^2 \mathbb{E}[Y_1^2]\frac{D_m}{n}.$$

Et enfin un choix optimal de modèle, si la régularité de h est connue :

Proposition 5. *Supposons que* $h = f \circ G^{-1}$ *soit un élément de l'espace de Besov* $\mathcal{B}_{2,\infty}^{\alpha}$*, avec* $\alpha > 0$ *fixé (*$\alpha \leq r+1$ *dans le cas où* S_m *est un espace de polynômes par morceaux de degrés inférieurs ou égaux à* r*). Alors, le choix d'un modèle* S_m *de dimension* D_m *de l'ordre de* $n^{1/(2\alpha+1)}$ *entraîne la majoration :*

$$\mathbb{E}\left[\|\hat{f}_m^{G,G} - f\|_g^2\right] \leq C n^{\frac{-2\alpha}{2\alpha+1}}.$$

De plus ce choix est optimal quant à la majoration du risque obtenue dans le Corollaire 4.

Preuve

Il s'agit de minimiser en $x \in \mathbb{R}_+^*$, la fonction $\chi(x) := C_1 x/n + C_2 x^{-2\alpha}$, où $C_1 = \phi_0^2\mathbb{E}[Y_1^2]$ et $C_2 = C_\alpha$. En dérivant, on obtient un unique point critique,

$$x_0 = (2\alpha C_2/C_1)^{1/(2\alpha+1)}\, n^{\frac{1}{2\alpha+1}},$$

et on vérifie qu'il s'agit bien du lieu d'un minimum, minimum dont la valeur est

$$\chi(x_0) = C n^{\frac{-2\alpha}{2\alpha+1}},$$

où C est une constante dépendant de C_1 et C_2.

□

6. Estimation adaptative

On ne suppose plus fixé un indice $m \in \mathcal{M}_n$. Disposant de la collection $(\hat{f}_m^{\hat{G}})_{m\in\mathcal{M}_n}$ d'estimateurs, on cherche à sélectionner un indice m convenable sur la base des observations.
On propose deux méthodes, l'une fondée sur un procédure classique de sélection de modèles par pénalisation, et l'autre fondée sur la comparaison des estimateurs ci-dessus deux à deux, basée sur des travaux de Lepski et Goldenshluger.

6.1. Construction de l'estimateur par méthode de Lepski.

On définit les deux quantités suivantes, quel que soit $m \in \mathcal{M}_n$, et $\delta > 0$:

$$\begin{aligned} &V(m) = 6 \times 6(1+2\delta)\phi_0^2\mathbb{E}\left[Y_1^2\right]\tfrac{D_m}{n}, \\ &A(m) = \max_{m'\in\mathcal{M}_n}\left(\left\|\hat{h}_{m'}^G - \hat{h}_{m\wedge m'}^G\right\|^2 - V(m')\right)_+, \end{aligned} \tag{5}$$

puisque l'on rappelle que la variance d'un estimateur $\hat{f}_m^{G,G}$ est majorée par $\phi_0^2\mathbb{E}[Y_1^2]D_m/n$.

$$\hat{m}^{G,l} = \operatorname{argmin}_{m\in\mathcal{M}_n}\{A(m) + V(m)\}. \tag{6}$$

L'estimateur sélectionné ainsi est donc $\hat{f}_{\hat{m}^{G,l}}^{G,G} = \hat{h}_{\hat{m}^{G,l}}^G \circ G$.

6.2. **Construction de l'estimateur par pénalisation.**

Soient $\theta > 2$ et $\delta > 0$ deux paramètres fixés. On définit l'application $\mathrm{pen}_{\theta,\delta}$:

$$m \in \mathcal{M}_n \quad \mapsto \quad \mathrm{pen}_{\theta,\delta}(m) := 6\theta(1+2\delta)\phi_0^2 \mathbb{E}[Y_1^2]\frac{D_m}{n}. \tag{7}$$

Soit aussi $\hat{m}^{G,p} \in \mathrm{argmin}_{m\in\mathcal{M}_n} \gamma_n^G(\hat{h}_m^G) + \mathrm{pen}_{\theta,\delta}(m)$.

L'estimateur sélectionné est $\hat{f}^{G,G}_{\hat{m}^{G,p}} = \hat{h}^G_{\hat{m}^{G,p}} \circ G$.

6.3. **Majoration du risque pour chacun des deux estimateurs.**

On va obtenir les deux théorèmes ci-dessous, sous les hypothèses suivantes :
– on se place sur l'un des modèles cités plus haut : trigonométriques, ondelettes à support compact, polynômes par morceaux dyadiques,
– le bruit ε_1 admet un moment d'ordre $2+p$, $p > 4$,
– la fonction h est bornée sur $[0;1]$.

Théorème 2. *L'estimateur sélectionné par méthode de Lepski vérifie l'inégalité-oracle suivante :*

$$\mathbb{E}\left[\left\|\hat{f}^{G,G}_{\hat{m}^{G,l}} - f\right\|_g^2\right] \leq C \min_{m\in\mathcal{M}_n}\left\{\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m}{n} + 15\left\|f_m^G - f\right\|_g^2\right\} + \frac{C}{n},$$

avec C une constante ne dépendant que de ϕ_0^2, $\|h\|_\infty$, et δ.

Théorème 3. *Soient $\theta > 2$ et $\delta > 0$ deux paramètres fixés. L'estimateur $\hat{f}^{G,G}_{\hat{m}^{G,p}}$ vérifie l'inégalité-oracle suivante :*

$$\mathbb{E}\left[\|\hat{f}^{G,G}_{\hat{m}^{G,p}} - f\|_g^2\right] \quad \leq \quad \min_{m\in\mathcal{M}_n}\left\{\frac{\theta+2}{\theta-2}\|f - f_m^G\|_g^2 + \frac{2\theta}{\theta-2}pen_{\theta,\delta}(m)\right\} + \frac{\theta^2}{\theta-2}\frac{C}{n}, \tag{8}$$

où l'on rappelle que $f_m^G = h_m \circ G$, h_m étant la projection orthogonale de h sur S_m, et où C est une constante ne dépendant que de ϕ_0^2, $\|f\|_\infty$, σ^2, $\mathbb{E}[|\varepsilon_1|^p]$, et δ.

Corollaire 6. *Sous les hypothèses des théorèmes précédents, et en supposant que la fonction h est dans l'espace de Besov $\mathcal{B}^\alpha_{2,\infty}$, pour $\alpha > 0$ ($\alpha \leq r+1$ dans le cas où S_m est un espace de polynômes par morceaux de degrés inférieurs ou égaux à r), et bornée sur $[0;1]$, pour $\hat{f}^{G,G}_{\hat{m}}$ l'un des deux estimateurs ci-dessus,*

$$\mathbb{E}\left[\|\hat{f}^{G,G}_{\hat{m}} - f\|_g^2\right] = O\left(n^{\frac{-2\alpha}{2\alpha+1}}\right).$$

Preuve du Corollaire 6

On reprend les inégalités oracles (par exemple (8) pour l'estimateur par pénalisation), on majore le biais par $C_\alpha D_m^{-2\alpha}$ en utilisant le lemme 12 de l'article déjà cité de Barron, Birgé et Massart, puis on calcule le minimum du membre de gauche de l'inégalité-oracle : celui-ci est atteint pour D_m de l'ordre de $n^{1/(2\alpha+1)}$, et vaut la vitesse $n^{-2\alpha/(2\alpha+1)} > n^{-1}$ annoncée.

□

Remarque : le jeu d'hypothèses suivant convient aussi, pour obtenir les mêmes résultats,
– on se place sur l'un des modèles cités plus haut : trigonométriques, ondelettes à support compact, polynômes par morceaux dyadiques.
– les dimensions des modèles vérifient $D_m \leq n^{1/2}$ (quel que soit $m \in \mathcal{M}_n$),

– le bruit ε_1 admet un moment d'ordre $2+p$, $p>2$,
– la fonction h est bornée sur $[0;1]$,
– (uniquement pour le corollaire, et à la place de l'hypothèse précédente) la fonction h est dans l'espace de Besov $\mathcal{B}^{\alpha}_{2,\infty}$, ($\alpha \leq r+1$ dans le cas où S_m est un espace de polynômes par morceaux de degrés inférieurs ou égaux à r), avec $\alpha > 1/2$ (ceci assure d'une part que D_m peut-être à la fois inférieur à $\sqrt{n}$ et de l'ordre de $n^{1/(2\alpha+1)}$; et d'autre part ceci assure automatiquement alors que h est bornée).

6.4. **Résultat préliminaire aux preuves des Théorèmes 2 et 3.**

Ces deux théorèmes sont fondés sur la majoration des déviations du processus empirique suivant : pour $t \in L^2([0;1])$,

$$\nu_n(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i t\circ G(X_i) - \langle t\circ G, f\rangle_g, \tag{9}$$

déviations étudiées par rapport à la quantité suivante, définie pour $m, m' \in \mathcal{M}_n$:

$$p(m,m') = 6(1+2\delta)\phi_0^2\mathbb{E}[Y_1^2]\frac{D_{m'\vee m}}{n}.$$

Lemme 7. *Avec les notations et hypothèses ci-dessus, quel que soit $m \in \mathcal{M}_n$,*

$$\mathbb{E}\left[\sum_{m'\in\mathcal{M}_n}\left(\sup_{\substack{t\in S_{m\vee m'}\\ \|t\|=1}} (\nu_n(t))^2 - p(m,m')\right)_+\right] \leq \frac{C}{n},$$

où C est une constante ne dépendant que de ϕ_0^2, $\|f\|_\infty$, $\mathbb{E}[f^2(X_1)]$, σ^2, $\mathbb{E}[|\varepsilon_1|^p]$, et δ.

6.5. **Preuve du Théorème 2.**

Dans toute la preuve C désignera une constante dépendant uniquement des quantités indiquées dans l'énoncé du théorème et pouvant varier d'une ligne à l'autre. On abrège $\hat{m}^{G,l}$ en $\hat{m}$, et $\hat{f}^{G,G,l}_{\hat{m}}$ en $\hat{f}^{G,G}_{\hat{m}}$.

On fixe $m \in \mathcal{M}_n$. On commence par décomposer

$$\begin{aligned}\left\|\hat{f}^{G,G}_{\hat{m}} - f\right\|_g^2 &= \left\|\hat{h}^G_{\hat{m}} - h\right\|^2,\\ &\leq 3\left\|\hat{h}^G_{\hat{m}} - \hat{h}^G_{m\wedge\hat{m}}\right\|^2 + 3\left\|\hat{h}^G_{m\wedge\hat{m}} - \hat{h}^G_{m}\right\|^2 + 3\left\|\hat{h}^G_{m} - h\right\|^2.\end{aligned}$$

En utilisant successivement la définition de A, puis celle de $\hat{m}$, on obtient,

$$\begin{aligned}\left\|\hat{f}^{G,G}_{\hat{m}} - f\right\|_g^2 &\leq 3\left(A(m)+V(\hat{m})\right) + 3\left(A(\hat{m})+V(m)\right) + 3\left\|\hat{h}^G_{m} - h\right\|^2,\\ &\leq 6\left(A(m)+V(m)\right) + 3\left\|\hat{h}^G_{m} - h\right\|^2.\end{aligned}$$

L'objectif est d'obtenir une inégalité de type oracle. Le terme $V(m)$ correspond au terme de variance et le terme $\left\|\hat{h}^G_{m} - h\right\|^2$ a déjà été étudié : il s'agit de la perte de l'estimateur $\hat{h}^G_m$, dont l'espérance est majorée par la somme d'un terme de variance de l'ordre de $V(m)$ et d'un terme de biais (au carré). On a donc

$$\mathbb{E}\left[\left\|\hat{h}^G_{\hat{m}} - h\right\|^2\right] \leq 6\mathbb{E}\left[A(m)\right] + 6V(m) + 3\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m}{n} + 3\left\|h_m - h\right\|^2.$$

Il reste à majorer le terme $A(m)$. On utilise le lemme suivant, ce qui conclut la preuve.

Lemme 8. *Sous les hypothèses du théorème, il existe une constante $C \geq 0$, telle que, pour tout $m \in \mathcal{M}_n$,*

$$\mathbb{E}\left[A(m)\right] \leq \frac{C}{n} + 12 \left\|h_m - h\right\|^2.$$

□

6.6. **Preuve du Théorème 3.**

Soit $m \in \mathcal{M}_n$ fixé. Ecrivons dans ce paragraphe $\gamma_n = \gamma_n^G$, pen $= \text{pen}_{\theta,\delta}$, $\hat{m} = \hat{m}^{G,p}$, $\hat{f}_{\hat{m}}^{G,G} = \hat{f}_{\hat{m}}^{G,G,p}$ pour alléger les notations.
On commence par utiliser la définition de $\hat{m}$ comme argument-minimum du contraste pénalisé :

$$\gamma_n(\hat{h}_{\hat{m}}^G) + \text{pen}(\hat{m}) \leq \gamma_n(\hat{h}_m^G) + \text{pen}(m).$$

On utilise ensuite celle de $\hat{h}_m^G$, comme minimiseur du contraste sur S_m, et le fait que $h_m \in S_m$.

$$\gamma_n(\hat{h}_m^G) \leq \gamma_n(h_m).$$

En combinant les deux inégalités, on a

$$\gamma_n(\hat{h}_m^G) - \gamma_n(h_m) \leq \text{pen}(m) - \text{pen}(\hat{m}). \tag{10}$$

On remplace ensuite le contraste par sa définition, en remarquant que pour tout $t \in L^2([0;1])$,

$$\|t\|^2 = \|t \circ G\|_g^2 = \|t \circ G - f\|_g^2 - \|f\|_g^2 + 2\langle t \circ G, f\rangle_g,$$

ce qui entraîne par l'inégalité (10),

$$\underbrace{\|\hat{h}_{\hat{m}}^G \circ G - f\|_g^2}_{\|\hat{f}_{\hat{m}}^{G,G} - f\|_g^2} \leq \underbrace{\|h_m \circ G - f\|_g^2}_{\|f_m^G - f\|_g^2} + \text{pen}(m) - \text{pen}(\hat{m}) + 2\nu_n(\hat{h}_{\hat{m}}^G - h_m), \tag{11}$$

où le processus ν_n a été défini ci-dessus (cf. l'equation (9)). Il s'agit ensuite de se débarrasser du caractère doublement aléatoire de $\nu_n(\hat{h}_{\hat{m}}^G - h_m)$: aléa du ν_n et aléa de son argument. Dans cet objectif on écrit :

$$\begin{aligned}
2\nu_n(\hat{h}_{\hat{m}}^G - h_m) &= 2\|\hat{h}_{\hat{m}}^G - h_m\|\nu_n\left(\frac{\hat{h}_{\hat{m}}^G - h_m}{\|\hat{h}_{\hat{m}}^G - h_m\|}\right) \text{ par linéarité de } \nu_n, \\
&\leq 2\|\hat{h}_{\hat{m}}^G - h_m\| \sup_{\substack{t \in S_m + S_{\hat{m}} \\ \|t\|=1}} |\nu_n(t)|, \\
&= 2\|\hat{h}_{\hat{m}}^G - h_m\| \sup_{t \in \mathcal{S}(m \vee \hat{m})} |\nu_n(t)| \text{ car les espaces sont emboîtés,} \\
&= 2\|\hat{f}_{\hat{m}}^{G,G} - f_m^G\|_g \sup_{t \in \mathcal{S}(m \vee \hat{m})} |\nu_n(t)|.
\end{aligned}$$

où l'on note ici et dans la suite $\mathcal{S}(p) = \{t \in S_p,\ \|t\| = 1\}$, pour $p \in \mathcal{M}_n$.
On introduit ensuite le paramètre $\theta > 0$, et on utilise l'inégalité $2xy \leq x^2/\theta + \theta y^2$, valable pour

tout couple de réels (x, y), pour en déduire :

$$\begin{aligned} 2\nu_n(\hat{h}^G_{\hat{m}} - h_m) &\leq \frac{\|\hat{f}^{G,G}_{\hat{m}} - f^G_m\|^2_g}{\theta} + \theta \sup_{t\in\mathcal{S}_{m\vee\hat{m}}} (\nu_n(t))^2, \\ &\leq \frac{2}{\theta}\left(\|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g + \|f^G_m - f\|^2_g\right) + \theta \sup_{t\in\mathcal{S}_{m\vee\hat{m}}} (\nu_n(t))^2. \end{aligned}$$

On injecte ceci dans l'inégalité (11) pour obtenir,

$$\left(1 - \frac{2}{\theta}\right) \|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g \leq \left(1 + \frac{2}{\theta}\right) \|f - f^G_m\|^2_g + \mathrm{pen}(m) - \mathrm{pen}(\hat{m}) + \theta \sup_{t\in\mathcal{S}(m\vee\hat{m})} (\nu_n(t))^2.$$

On introduit maintenant $p(m, \hat{m})$, quantité positive définie aussi ci-dessus, et l'on écrit, pour $\theta > 2$,

$$\begin{aligned} \|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g &\leq \frac{\theta}{\theta-2}\left\{\frac{\theta+2}{\theta}\|f - f^G_m\|^2_g + \mathrm{pen}(m) + \right. \\ &\left. \theta\left(\sup_{t\in\mathcal{S}_{m\vee\hat{m}}} (\nu_n(t))^2 - p(m,\hat{m})\right)_+ + \theta p(m,\hat{m}) - \mathrm{pen}(\hat{m})\right\}. \end{aligned}$$

Il reste un double aléa dans le "sup" qui apparaît dans la majoration : aléa dans l'indice du modèle en plus de l'aléa présent dans la définition de ν_n. Pour s'en débarrasser, on majore brutalement de la façon suivante :

$$\begin{aligned} \|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g &\leq \frac{\theta+2}{\theta-2}\|f - f^G_m\|^2_g + \frac{\theta}{\theta-2}\mathrm{pen}(m) \\ &+ \frac{\theta^2}{\theta-2}\sum_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m\vee\hat{m})} (\nu_n(t))^2 - p(m,m')\right)_+ \\ &+ \frac{\theta}{\theta-2}\left(\theta p(m,\hat{m}) - \mathrm{pen}(\hat{m})\right). \end{aligned} \tag{12}$$

La clé de la preuve consiste à utiliser le Lemme 7 ; on obtient alors, en partant de l'équation (12),

$$\mathbb{E}\left[\|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g\right] \leq \frac{\theta+2}{\theta-2}\|f - f^G_m\|^2_g + \frac{\theta}{\theta-2}\mathrm{pen}(m) + \frac{\theta^2}{\theta-2}\frac{C}{n} + \frac{\theta}{\theta-2}\mathbb{E}\left[\theta p(m,\hat{m}) - \mathrm{pen}(\hat{m})\right].$$

La définition de $\mathrm{pen}(m)$ impose ensuite que

$$\mathbb{E}\left[\theta p(m,\hat{m}) - \mathrm{pen}(\hat{m})\right] \leq \mathrm{pen}(m).$$

puisque l'on a

$$p(m,\hat{m}) \leq 6(1+2\delta)\phi_0^2\mathbb{E}[Y_1^2]\frac{D_m + D_{\hat{m}}}{n},$$

Ceci entraîne,

$$\mathbb{E}\left[\|\hat{f}^{G,G}_{\hat{m}} - f\|^2_g\right] \leq \frac{\theta+2}{\theta-2}\|f - f^G_m\|^2_g + \frac{2\theta}{\theta-2}\mathrm{pen}(m) + \frac{\theta^2}{\theta-2}\frac{C}{n},$$

ce qui est le résultat annoncé.

□

6.7. **Preuve du Lemme 7.**

Pour plus de lisibilité on décompose la preuve en 6 étapes.

• **Première étape : décomposition de ν_n**
L'idée directrice est l'application de l'inégalité de Talagrand à

$$\nu_n(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ G(X_i) - \langle t \circ G, f\rangle_g,$$

pour $t \in L^2([0;1])$. Mais son utilisation nécessite des quantités bornées : or, $Y_i t \circ G(X_i)$ ne l'est pas a priori. On écrit donc,

$$\nu_n(t) = \nu_n^{(1)}(t) + \nu_n^{(2,1)}(t) + \nu_n^{(2,2)}(t),$$

avec

$$\nu_n^{(1)}(t) = \frac{1}{n}\sum_{i=1}^{n} f(X_i) t \circ G(X_i) - \langle t \circ G, f\rangle_g,$$

$$\nu_n^{(2,1)}(t) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{1}_{|\varepsilon_i|\le\kappa_n} t \circ G(X_i) - \mathbb{E}\left[\varepsilon_i \mathbf{1}_{|\varepsilon_i|\le\kappa_n} t \circ G(X_i)\right],$$
$$\nu_n^{(2,2)}(t) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{1}_{|\varepsilon_i|>\kappa_n} t \circ G(X_i) + \mathbb{E}\left[\varepsilon_i \mathbf{1}_{|\varepsilon_i|\le\kappa_n} t \circ G(X_i)\right],$$

où

$$\kappa_n = c\frac{\sqrt{n}}{\ln(n)},$$

avec c une constante dépendant du modèle choisi (trigonométrique, ondelettes, polynômes par morceaux).
Le dernier terme s'écrit aussi :

$$\nu_n^{(2,2)}(t) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{1}_{|\varepsilon_i|>\kappa_n} t \circ G(X_i) - \mathbb{E}\left[\varepsilon_i \mathbf{1}_{|\varepsilon_i|>\kappa_n} t \circ G(X_i)\right],$$

en utilisant le fait que $\mathbb{E}(\varepsilon_i) = 0$ et l'indépendance de X_i d'avec ε_i.
On a ainsi

$$\nu_n(t)^2 \le 3\left(\left(\nu_n^{(1)}(t)\right)^2 + \left(\nu_n^{(2,1)}(t)\right)^2 + \left(\nu_n^{(2,2)}\right)^2\right),$$

et si on décompose $p(m,\hat{m}) = p_1(m,\hat{m}) + p_2(m,\hat{m})$, on a, quel que soit $m' \in \mathcal{M}_n$

$$\begin{aligned}\left(\sup_{t\in\mathcal{S}(m'\vee m)} \nu_n(t)^2 - p(m,m')\right)_+ \quad &\le \quad 3\left\{\left(\sup_{t\in\mathcal{S}(m'\vee m)} \left(\nu_n^{(1)}(t)\right)^2 - \frac{p_1(m,m')}{3}\right)_+\right. \\ &+ \left(\sup_{t\in\mathcal{S}(m'\vee m)} \left(\nu_n^{(2,1)}(t)\right)^2 - \frac{p_2(m,m')}{3}\right)_+ \\ &\left. + \sup_{t\in\mathcal{S}(m'\vee m}) \left(\nu_n^{(2,2)}(t)\right)^2\right\}\end{aligned}$$

On va maintenant majorer chacun des trois termes, en appliquant l'inégalité de Talagrand pour les deux premiers, et une majoration de type Markov pour le suivant.

• **Deuxième étape : Majoration du premier terme**
Pour $t \in \mathcal{S}(m' \vee m)$, on s'intéresse à

$$\begin{aligned} \nu_n^{(1)}(t) &= \frac{1}{n}\sum_{i=1}^n f(X_i)t\circ G(X_i) - \langle t\circ G, f\rangle_g, \\ &= \frac{1}{n}\sum_{i=1}^n f_t(X_i) - \mathbb{E}[f_t(X_i)], \end{aligned}$$

où $f_t : x \mapsto f(x)t\circ G(x)$.
Pour appliquer l'inégalité de Talagrand (1), on commence par déterminer les quantités $M_1^{(1)}$, $v^{(1)}$, $H^{(1)}$ qui interviennent.

- $M_1^{(1)}$ est un majorant uniforme sur $\mathcal{S}(m' \vee m)$ des $\|f_t\|_\infty$. Or, $\|f_t\|_\infty \leq \|f\|_\infty \|t\|_\infty$, et, comme $t \in S_{m\vee m'}$, il existe $(b_j)_{j=1,\dots,D_{m\vee m'}}$ une famille de réels telle que $t = \sum b_j\varphi_j$. Comme de plus $\|t\| = 1$, on a $\sum b_j^2 = 1$. Donc l'inégalité de Cauchy-Schwarz entraîne que $\|t\|_\infty^2 \leq \sum \|\varphi_j\|_\infty^2 \leq \phi_0^2 D_m$. Donc on peut choisir
$$M_1^{(1)} = \phi_0 \|f\|_\infty \sqrt{D_{m\vee m'}}.$$
- On cherche $H^{(1)}$ majorant $\mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)} \nu_n^{(1)}(t)\right]$. Par l'inégalité de Cauchy-Schwarz, il suffit de majorer
$$\mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)} \left(\nu_n^{(1)}(t)\right)^2\right],$$
et en écrivant que $t \in \mathcal{T}_{m'}$ comme ci-dessus, on obtient,
$$\sup_{t\in\mathcal{S}(m'\vee m)} \left(\nu_n^{(1)}(t)\right)^2 \leq \sum_{j=1}^{D_{m\vee m'}} \left(\nu_n^{(1)}(\varphi_j)\right)^2.$$
De plus,
$$\begin{aligned} \mathbb{E}\left[\left(\nu_n^{(1)}(\varphi_j)\right)^2\right] &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n f(X_i)\varphi_j(G(X_i))\right), \\ &= \frac{1}{n}\mathrm{Var}(f(X_1)\varphi_j(G(X_1))), \\ &\leq \frac{1}{n}\mathbb{E}\left[f^2(X_1)\varphi_j^2(G(X_1))\right], \end{aligned}$$
donc,
$$\mathbb{E}\left[\sum_{j=1}^{D_{m\vee m'}} \left(\nu_n^{(1)}(\varphi_j)\right)^2\right] \leq \frac{1}{n}\mathbb{E}\left[f^2(X_1)\right]\phi_0^2 D_{m\vee m'},$$
et on peut donc choisir :
$$\left(H^{(1)}\right)^2 = \frac{D_{m\vee m'}}{n}\mathbb{E}\left[f^2(X_1)\right]\phi_0^2.$$

– $v^{(1)}$ est défini comme un majorant de $\sup_{t\in\mathcal{S}(m'\vee m)} 1/n\sum_i \ \mathrm{Var}(f_t(X_i))$. Comme les données sont i.i.d, cette quantité est égale à

$$\begin{aligned}\mathrm{Var} f_t(X_1) &\leq \mathbb{E}\left[f_t(X_1)^2\right] = \int_a^b t^2(G(x))f^2(x)g(x)dx,\\ &= \int_0^1 t^2(u)(f\circ G^{-1})^2(u)du \leq \|f\|_\infty^2 \int_0^1 t^2(u)du,\\ &= \|f\|_\infty^2 := v^{(1)}.\end{aligned}$$

On applique maintenant l'inégalité :

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(1)}(t)\right)^2 - 2(1+2\delta)(H^{(1)})^2\right)_+\right]$$
$$\leq \frac{4}{K_1}\left\{\frac{v^{(1)}}{n}\exp\left(-K_1\delta\frac{n\left(H^{(1)}\right)^2}{v^{(1)}}\right) + \frac{49}{K_1C^2(\delta)}\frac{\left(M_1^{(1)}\right)^2}{n^2}\exp\left(\frac{-\sqrt{2}K_1C(\delta)\sqrt{\delta}}{7}\frac{nH^{(1)}}{M_1^{(1)}}\right)\right\},$$

où $C(\delta) = (\sqrt{1+\delta}-1)\wedge 1$ et $K_1 = 1/6$. En remplaçant ensuite $M_1^{(1)}, v^{(1)}, H^{(1)}$ par leurs valeurs ci-dessus, et en choisissant

$$p_1(m,m') := 3\times 2(1+2\delta)\left(H^{(1)}\right)^2,$$

il vient, en sommant sur $m'\in\mathcal{M}_n$,

$$\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(1)}(t)\right)^2 - \frac{p_1(m,m')}{3}\right)_+\right]$$
$$\leq \frac{4}{K_1}\|f\|_\infty^2\left\{\frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\bar{k}D_{m\vee m'}\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{1}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m\vee m'}\exp\left(-\bar{\bar{k}}\sqrt{n}\right)\right\}.$$

où l'on a noté,

$$\begin{cases}\bar{k} = \dfrac{k_1\delta\phi_0^2\mathbb{E}[f^2(X_1)]}{\|f\|_\infty^2},\\[2ex] \bar{\bar{k}} = \dfrac{\sqrt{2}K_1C(\delta)\sqrt{\delta}\sqrt{\mathbb{E}[f^2(X_1)]}}{7\|f\|_\infty}.\end{cases}$$

Le membre de droite de l'inégalité ci dessus fait apparaître deux sommes du même type que l'on majore de la façon suivante :

$$\begin{aligned}\sum_{m'\in\mathcal{M}_n}\chi(m\vee m') &\leq \mathrm{Card}\left\{m'\in\mathcal{M}_n,\, m'\leq m\right\}\chi(m) + \sum_{m'\in\mathcal{M}_n}\chi(m'),\\ &\leq m\chi(m) + \sum_{m'\in\mathcal{M}_n}\chi(m'),\end{aligned}$$

ce qui permet d'obtenir,

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(1)}(t)\right)^2 - \frac{p_1(m,m')}{3}\right)_+\right] \leq T_1,$$

où

$$T_1 = \frac{4}{K_1}\|f\|_\infty^2 \left\{ \frac{m}{n}\exp\left(-\bar{k}D_m\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{m}{n^2}D_m\exp\left(-\bar{\bar{k}}\sqrt{n}\right) \right.$$
$$\left. + \frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\bar{k}D_{m'}\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{\exp\left(-\bar{\bar{k}}\sqrt{n}\right)}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m'}\right\}.$$

• **Troisième étape : Majoration du deuxième terme**
Pour $t \in \mathcal{S}_{m\vee m'}$, on s'intéresse à

$$\begin{aligned} \nu_n^{(2,1)}(t) &= \frac{1}{n}\sum_{i=1}^n \varepsilon_i \mathbf{1}_{|\varepsilon_i|\le\kappa_n} t\circ G(X_i) - \mathbb{E}\left[\varepsilon_i\mathbf{1}_{|\varepsilon_i|\le\kappa_n} t\circ G(X_i)\right], \\ &= \frac{1}{n}\sum_{i=1}^n g_t(\varepsilon_i, X_i) - \mathbb{E}\left[g_t(\varepsilon_i, X_i)\right], \end{aligned}$$

où $g_t(\varepsilon, x) := \varepsilon\mathbf{1}_{|\varepsilon|\le\kappa_n} t\circ G(x)$.
Comme pour le premier terme, pour appliquer l'inégalité de Talagrand, on commence par déterminer les quantités $M_1^{(2)}, v^{(2)}, H^{(2)}$, définit ci-dessus (en remplaçant f_t par g_t) qui interviennent :

– On écrit $\|g_t\|_\infty \le \kappa_n\|t\|_\infty$, ce qui permet de prendre, en suivant le choix du $M_1^{(1)}$ ci-dessus,
$$M_1^{(2)} = \kappa_n\phi_0\sqrt{D_{m\vee m'}}.$$
– De la même façon, on écrit aussi
$$\begin{aligned} \mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,1)}(t)\right)^2\right] &\le \frac{1}{n}\sum_{j=1}^{D_{m\vee m'}}\mathbb{E}\left[\varepsilon_1^2\mathbf{1}_{|\varepsilon_1|\le\kappa_n}\varphi_j^2(G(X_1))\right], \\ &\le \frac{1}{n}\sum_{j=1}^{D_{m\vee m'}}\mathbb{E}\left[\varepsilon_1^2\varphi_j^2(G(X_1))\right], \\ &\le \phi_0^2\sigma^2\frac{D_{m\vee m'}}{n} := \left(H^{(2)}\right)^2 \end{aligned}$$
– Pour la quantité $v^{(2)}$, comme ci-dessus, la variance de la somme est n fois la variance individuelle, puis on majore la variance individuelle par l'espérance du carré, et l'indicatrice par 1 pour faire sortir $\sigma^2\mathbb{E}(t\circ G(X_1)^2) = \sigma^2 := v^{(2)}$.

Comme pour le premier terme, on applique l'inégalité, on somme sur m', et en choisissant
$$p_2(m,m') := 3\times 2(1+2\delta)\left(H^{(2)}\right)^2,$$
il vient,
$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,1)}(t)\right)^2 - \frac{p_2(m,m')}{3}\right)_+\right]$$
$$\le \frac{4}{K_1}\left\{\sigma^2\frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\tilde{k}D_{m\vee m'}\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{\kappa_n}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m\vee m'}\exp\left(-\tilde{\tilde{k}}\sqrt{n}\right)\right\}.$$

où l'on a noté,

$$\left\{\begin{array}{l}\tilde{k} = k_1\delta\phi_0^2\mathbb{E},\\[2mm] \tilde{\tilde{k}} = \dfrac{\sqrt{2}K_1C(\delta)\sqrt{\delta}\sigma}{7}.\end{array}\right.$$

On conclut encore comme pour le premier terme :

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,1)}(t)\right)^2 - \frac{p_2(m,m')}{3}\right)_+\right] \leq T_2,$$

où

$$\begin{aligned} T_2 &= \frac{4}{K_1}\left\{\sigma^2\frac{m}{n}\exp\left(-\tilde{k}D_m\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{m}{n^2}\kappa_n^2 D_m\exp\left(-\tilde{\tilde{k}}\frac{\sqrt{n}}{\kappa_n}\right)\right.\\ &\left.+\sigma^2\frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\tilde{k}D_{m'}\right) + \frac{49\phi_0^2}{K_1C^2(\delta)}\frac{\exp\left(-\tilde{\tilde{k}}\frac{\sqrt{n}}{\kappa_n}\right)\kappa_n^2}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m'}\right\}.\end{aligned}$$

• **Quatrième étape : Majoration du troisième terme**

Pour $t\in\mathcal{S}(m'\vee m)$, on s'intéresse à

$$\nu_n^{(2,2)}(t) = \frac{1}{n}\sum_{i=1}^n \varepsilon_i\mathbf{1}_{|\varepsilon_i|>\kappa_n}t\circ G(X_i) - \mathbb{E}\left[\varepsilon_i\mathbf{1}_{|\varepsilon_i|>\kappa_n}t\circ G(X_i)\right].$$

Ce processus empirique n'est pas borné, donc on ne peut espérer lui appliquer l'inégalité de Talagrand. En écrivant ce que signifie $t\in\mathcal{T}_{m'}$, comme explicité ci-dessus dans l'étude du premier terme, on obtient :

$$\mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,2)}(t)\right)^2\right] \leq \sum_{j=1}^{D_{m\vee m'}}\mathbb{E}\left[\left(\nu_n^{(2,2)}(\varphi_j)^2\right)\right],$$

et

$$\begin{aligned}\mathbb{E}\left[\left(\nu_n^{(2,2)}(\varphi_j)^2\right)\right] &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n\varepsilon_i\mathbf{1}_{|\varepsilon_i|>\kappa_n}\varphi_j\circ G(X_i)\right),\\ &\leq \frac{1}{n}\mathbb{E}\left[\varepsilon_1^2\mathbf{1}_{|\varepsilon_1|>\kappa_n}(\varphi_j\circ G(X_1))^2\right],\end{aligned}$$

D'où

$$\begin{aligned}\sum_{j=1}^{D_{m\vee m'}}\mathbb{E}\left[\left(\nu_n^{(2,2)}(\varphi_j)^2\right)\right] &\leq \frac{1}{n}\mathbb{E}\left[\varepsilon_1^2\mathbf{1}_{|\varepsilon_1|>\kappa_n}\right]\phi_0^2D_{m\vee m'},\\ &\leq \frac{1}{n}\phi_0^2\kappa_n^{-p}\mathbb{E}\left[|\varepsilon_1|^{2+p}\mathbf{1}_{|\varepsilon_1|>\kappa_n}\right]D_{m\vee m'},\\ &\leq \frac{1}{n}\phi_0^2\kappa_n^{-p}\mathbb{E}\left[|\varepsilon_1|^{2+p}\right]D_{m\vee m'},\end{aligned}$$

On en déduit donc :

$$\mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,2)}(t)\right)^2\right] \leq \phi_0^2\mathbb{E}\left[|\varepsilon_1|^{2+p}\right]\frac{D_{m\vee m'}\kappa_n^{-p}}{n}.$$

En sommant sur m' et en appliquant la même démarche que pour les deux premiers termes, on obtient :

$$\mathbb{E}\left[\sup_{t\in\mathcal{S}(m'\vee m)}\left(\nu_n^{(2,2)}(t)\right)^2\right]\leq T_3,$$

où

$$T_3=\phi_0^2\mathbb{E}\left[|\varepsilon_1|^{2+p}\right]\left(\frac{mD_m\kappa_n^{-p}}{n}+\frac{\kappa_n^{-p}}{n}\sum_{m'\in\mathcal{M}_n}D'_m\right).$$

En résumé, nous avons obtenu :

$$\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m'\vee m)}\nu_n(t)^2-p(m,m')\right)_+\right]\leq 3(T_1+T_2+T_3)$$

où

$$T_1=\frac{4}{K_1}\|f\|_\infty^2\left\{\frac{m}{n}\exp\left(-\bar{k}D_m\right)+\frac{49\phi_0^2}{K_1C^2(\delta)}\frac{m}{n^2}D_m\exp\left(-\bar{\bar{k}}\sqrt{n}\right)\right.$$

$$\left.+\frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\bar{k}D_{m'}\right)+\frac{49\phi_0^2}{K_1C^2(\delta)}\frac{\exp\left(-\bar{\bar{k}}\sqrt{n}\right)}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m'}\right\},$$

$$T_2=\frac{4}{K_1}\left\{\sigma^2\frac{m}{n}\exp\left(-\tilde{k}D_m\right)+\frac{49\phi_0^2}{K_1C^2(\delta)}\frac{m}{n^2}\kappa_n^2D_m\exp\left(-\tilde{\tilde{k}}\frac{\sqrt{n}}{\kappa_n}\right)\right.$$

$$\left.+\sigma^2\frac{1}{n}\sum_{m'\in\mathcal{M}_n}\exp\left(-\tilde{k}D_{m'}\right)+\frac{49\phi_0^2}{K_1C^2(\delta)}\frac{\exp\left(-\tilde{\tilde{k}}\frac{\sqrt{n}}{\kappa_n}\right)\kappa_n^2}{n^2}\sum_{m'\in\mathcal{M}_n}D_{m'}\right\},$$

$$T_3=\phi_0^2\mathbb{E}\left[|\varepsilon_1|^{2+p}\right]\left(\frac{mD_m\kappa_n^{-p}}{n}+\frac{\kappa_n^{-p}}{n}\sum_{m'\in\mathcal{M}_n}D'_m\right).$$

• **Cinquième étape : Conclusion de la majoration du processus empirique dans le cas des modèles fondés sur la base trigonométrique**

On choisit maintenant le modèle fondé sur la base de Fourier pour conclure la majoration. Dans ce cadre, $D_m=2m+1$, $m\leq[n/2]-1$, donc $2m+1\leq n$, et $m\leq n$. On exploite ceci pour majorer les quantités T_1,T_2,T_3.
On commence par considérer séparément les termes présents dans T_1,T_2,T_3, et faisant intervenir m et n. Dans les calculs qui suivent, $\bar{k}$ pourra être remplacé par $\tilde{k}$ quand on majorera T_2 au lieu de T_1.

$$\frac{m}{n}\exp(-\bar{k}D_m)=\exp(-\bar{k})\frac{m}{n}\exp(-2\bar{k}m)\leq\frac{\exp(-\bar{k}-1)\frac{1}{2\bar{k}}}{n},$$

la dernière majoration découlant de l'étude de la fonction $u\mapsto u\exp(-2\bar{k}u)$ dont le maximum est atteint en $1/2\bar{k}$.
Ensuite,

$$\frac{mD_m}{n^2}\leq\frac{n([n/2]-1)}{n^2}\leq 1/2.$$

Puis,

$$\sum_{m' \in \mathcal{M}_n} \exp(-\bar{k} D'_m) \leq \sum_{m=1}^{\infty} \exp(-\bar{k} D'_m) = \frac{\exp(-\bar{k})}{1 - \exp(-2\bar{k})},$$

et

$$\begin{aligned} \sum_{m' \in \mathcal{M}_n} D'_m &\leq 2 \sum_{m'=1}^{[n/2]-1} m + \mathrm{Card}(\mathcal{M}_n), \\ &\leq [n/2]([n/2]-1) + [n/2] - 1, \\ &\leq \frac{n^2}{4} - 1 \leq n^2/4, \end{aligned}$$

d'où

$$\frac{\exp(-\bar{\bar{k}}\sqrt{n})}{n^2} \sum_{m' \in \mathcal{M}_n} D'_m \leq (1/4) \exp(-\bar{\bar{k}}\sqrt{n}) \leq \frac{e^{-2}}{\bar{\bar{k}}^2} \frac{1}{n},$$

comme le montre l'étude de $u \mapsto u \exp(-\bar{\bar{k}}\sqrt{u})$, qui atteint son maximum en $4/\bar{\bar{k}}^2$. Ceci nous permet déjà d'écrire, en notant que $\phi_0^2 = 2$ ici,

$$T_1 \leq \frac{C_1}{n}$$

où C_1 est une constante dépendant de δ, $\|f\|_\infty$, $\mathbb{E}[f^2(X_1)]$, mais ni de n ni de m : précisément,

$$C_1 = \frac{4\|f\|_\infty^2}{K_1} \left\{ \frac{e^{-\bar{k}-1}}{2\bar{k}} + \frac{196 e^{-2}}{K_1 C^2(\delta) \bar{k}^2} + \frac{e^{-\bar{k}}}{1 - e^{-2\bar{k}}} + \frac{98 e^{-2}}{K_1 C^2(\delta) \bar{\bar{k}}^2} \right\}.$$

Il s'agit ensuite de montrer que T_2 est en $O(1/n)$.
On rappelle que $\kappa_n \leq n$. Ceci, ainsi que l'utilisation des majorations ci dessus, montrent que l'on a

$$\frac{\kappa_n}{n^2} \exp\left(-\frac{\tilde{\bar{\bar{k}}}\sqrt{n}}{\kappa_n}\right) \sum_{m' \in \mathcal{M}_n} D'_m \leq \frac{1}{4} n \exp\left(-\frac{\tilde{\bar{\bar{k}}}\sqrt{n}}{\kappa_n}\right).$$

On impose que cette dernière quantité soit majorer par C/n où C est une constante. Choisissons par exemple $C = 1/4$ pour fixer les idées. On obtient donc

$$\exp\left(-\frac{\tilde{\bar{\bar{k}}}\sqrt{n}}{\kappa_n}\right) \leq \frac{1}{n^2} \implies \kappa_n \leq \frac{\tilde{\bar{\bar{k}}}}{2} \frac{\sqrt{n}}{\ln(n)}.$$

Le choix de $\kappa_n = \frac{\tilde{\bar{\bar{k}}}}{2} \frac{\sqrt{n}}{\ln(n)}$ permet d'écrire,

$$\begin{aligned} m D_m \frac{\kappa_n^2}{n^2} \exp\left(-\frac{\tilde{\bar{\bar{k}}}\sqrt{n}}{\kappa_n}\right) &= m D_m \frac{\tilde{\bar{\bar{k}}}}{4} \frac{n}{n^2 \ln^2(n)} \exp(-2\ln(n)), \\ &\leq \frac{\tilde{\bar{\bar{k}}}}{4} n^2 \frac{n}{n^2 \ln^2(n)} n^{-2} = n^{-1} (\ln(n))^{-2} \frac{\tilde{\bar{\bar{k}}}}{4} < \frac{\tilde{\bar{\bar{k}}}}{4n}. \end{aligned}$$

Ainsi,

$$T_2 \leq \frac{C_2}{n}$$

où C_2 est une constante dépendant de δ, ϕ_0^2, σ^2, mais ni de n ni de m : précisément,

$$C_2 = \frac{4}{K_1}\left\{\frac{e^{-\tilde{k}-1}}{2\tilde{k}}\sigma^2 + \frac{49\tilde{\tilde{k}}^2}{2K_1C^2(\delta)} + \sigma^2\frac{e^{-\tilde{k}}}{1-e^{-2\tilde{k}}} + \frac{49}{2K_1C^2(\delta)}\right\}.$$

Enfin, il s'agit de montrer que l'on a $T_3 = O(1/n)$, pour le choix de la puissance d'intégrabilité de ε_1 : $2+p$ avec $p > 4$. On a

$$mD_m\frac{\kappa_n^{-p}}{n} \le \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p} n\frac{n^{-p/2}}{(\ln(n))^{-p}} = \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p} n^{1-p/2}(\ln(n))^p,$$

et d'après les calculs qui précèdent

$$\frac{\kappa_n^{-p}}{n}\sum_{m'\in\mathcal{M}_n} D'_m \le \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p}\frac{\kappa_n^{-p}}{n}\frac{1}{4}n^2 = \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p}\frac{1}{4}n^{1-p/2}(\ln(n))^p.$$

Ainsi, les deux termes intervenant dans la majoration de T_3 sont du même ordre de grandeur. Comme $p > 4$,

$$1-\frac{p}{2} < -1.$$

Par suite,

$$T_3 \le \frac{C_3}{n},$$

où $C_3 = \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p}\frac{5}{4}\phi_0^2\mathbb{E}[\varepsilon_1^p] = \left(\frac{\tilde{\tilde{k}}}{2}\right)^{-p}\frac{5}{2}\mathbb{E}[\varepsilon_1^p]$.

Remarquons que c'est à ce niveau que le choix $p = 4$ entraîne un perte d'un facteur $\ln(n)^4$ dans la majoration : il reste alors

$$T_3 \le \frac{C_3\ln^4(n)}{n},$$

ce qui suffit tout de même pour avoir une majoration négligeable de l'espérance du processus empirique (cf. les remarques suivant le théorème).

De plus, notons aussi que si l'on suppose $D_m \le \sqrt{n}$, l'ordre de grandeur des deux termes devient $n^{-p/2}\ln(n)$, et l'hypothèse d'intégrabilité moins lourde $p > 2$ entraîne $-p/2 < -1$, et le même résultat sur T_3.

Finalement, on obtient,

$$\mathbb{E}\left[\sum_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m\vee m')}(\nu_n(t))^2 - p(m,m')\right)_+\right] \le \frac{C}{n},$$

où $C = 3(C_1+C_2+C_3)$, et $p(m,m') = (p_1+p_2)(m,m') = 6(1+2\delta)\phi_0^2(\mathbb{E}[f^2(X_1)]+\sigma^2)\frac{D_{m\vee m'}}{n} = 6(1+2\delta)\phi_0^2\mathbb{E}[Y_1^2]\frac{D_{m\vee m'}}{n}$, ce qui conclut la preuve dans le cas des modèles trigonométriques.

• Sixième étape : Conclusion de la majoration du processus empirique dans le cas des modèles fondés sur les polynômes par morceaux dyadiques ou les ondelettes

On rappelle que dans ce cadre, $D_m = k2^m$ (où $k = (r+1)$ pour les polynômes par morceaux, $k = 2$ pour les ondelettes), et la condition $D_m \le n$ entraîne

$$m \le \left[\frac{\ln(n)-\ln(k)}{\ln(2)}\right] = m_{\max}.$$

On reprend la démarche adoptée au point précédent, c'est à dire que l'on majore les quantités dépendants de n et m et intervenant dans T_1, T_2 ou T_3. On sera cependant moins précis sur les constantes présentes dans les majorations, la définition même de $D_m = 2^m$ compliquant quelque peu leurs expressions explicites (par rapport au $D_m = 2m+1$ précédent). Ainsi, dans les lignes qui suivent, c désigne une constante qui peut changer d'une ligne à l'autre.
L'étude de la fonction $u \mapsto u\exp(-k\bar{k}2^u)$ (atteignant un maximum en $u_0 = -\ln(k\bar{k}\ln(2))/\ln(2)$) nous permet d'affirmer :

$$m\exp\left(-\bar{k}D_m\right) \leq c.$$

Ensuite,

$$\frac{mD_m}{n^2} \leq \frac{n\frac{\ln(n)}{\ln(2)}}{n^2} = c\frac{\ln(n)}{n} \leq c.$$

Puis, $\exp(-\bar{k}D_{m'})$ est le terme général d'une série convergente (somme sur $m' \in \mathcal{M}_n \subset \mathbb{N}$), et enfin,

$$\sum_{m'\in\mathcal{M}_n} D_{m'} = k\sum_{m'=1}^{m_{\max}} 2^{m'} = k\left(2^{m_{\max}+1}-2\right) \leq cn,$$

par construction de $m_{\max}$.
Ainsi, aux constantes différentes près, on est exactement dans le même cas que dans le cas du modèle fondé sur la base trigonométrique, et tout le raisonnememnt fait précédemment s'applique. Finalement, on obtient aussi,

$$\mathbb{E}\left[\sum_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m\vee m')}(\nu_n(t))^2 - p(m,m')\right)_+\right] \leq \frac{C}{n},$$

où C est une constante dépendant de ϕ_0^2, $\|f\|_\infty$, $\mathbb{E}[f^2(X_1)]$, σ^2, $\mathbb{E}[|\varepsilon_1|^p]$, et δ.

□

6.8. **Preuve du Lemme 8.**
On doit majorer, pour chaque $m \in \mathcal{M}_n$,

$$A(m) = \max_{m'\in\mathcal{M}_n}\left\{\left\|\hat{h}^G_{m'} - \hat{h}^G_{m\wedge m'}\right\|^2 - V(m)\right\}_+.$$

Pour chaque indice $m' \in \mathcal{M}_n$, on décompose,

$$\left\|\hat{h}^G_{m'} - \hat{h}^G_{m\wedge m'}\right\|^2 \leq 3\left\|\hat{h}^G_{m'} - h_{m'}\right\|^2 + 3\left\|h_{m'} - h_{m\wedge m'}\right\|^2 + 3\left\|h_{m\wedge m'} - \hat{h}^G_{m\wedge m'}\right\|^2.$$

Ainsi,

$$\begin{aligned} A(m) &\leq 3\max_{m'\in\mathcal{M}_n}\left[\left\|\hat{h}^G_{m'} - h_{m'}\right\|^2 - \frac{V(m')}{6}\right]_+ + 3\max_{m'\in\mathcal{M}_n}\left[\left\|h_{m\wedge m'} - \hat{h}^G_{m\wedge m'}\right\|^2 - \frac{V(m')}{6}\right]_+ \\ &\quad +3\max_{m'\in\mathcal{M}_n}\left\|h_{m'} - h_{m\wedge m'}\right\|^2, \\ &:= 3\left(T_a + T_b^m + T_c^m\right). \end{aligned}$$

On a donc trois termes dont on doit majorer l'espérance.

6.8.1. *Majoration de T_a.*

On majore brutalement le premier de la façon suivante :

$$\mathbb{E}\left[T_a\right] \leq \sum_{m' \in \mathcal{M}_n} \mathbb{E}\left[\left\{\left\|\hat{h}^G_{m'} - h_{m'}\right\|^2 - \frac{V(m')}{6}\right\}_+\right].$$

Et l'on remarque que l'on a

$$\begin{aligned}\left\|\hat{h}^G_{m'} - h_{m'}\right\|^2 &= \sum_{j=1}^{D_{m'}} \left(\hat{a}^G_j - a_j\right)^2, \\ &= \sum_{j=1}^{D_{m'}} \left(\frac{1}{n}\sum_{i=1}^{n}\left(Y_i\varphi_j \circ G(X_i) - \mathbb{E}\left[Y_i\varphi_j \circ G(X_i)\right]\right)\right)^2, \\ &= \sum_{j=1}^{D_{m'}} \nu_n^2(\varphi_j),\end{aligned}$$

où ν_n est le processus empirique centré défini par l'égalité (9). De plus, on a l'égalité $\sum_{j=1}^{D_{m'}} \nu_n^2(\varphi_j) = \sup_{t\in\mathcal{S}(m')} \nu_n^2(t)$, où $\mathcal{S}(m')$ désigne la sphère unité du modèle $S_{m'}$. En effet, si $t \in \mathcal{S}(m')$, $t = \sum_{j=1}^{D_{m'}} b_j\varphi_j$, avec $\sum_j b_j^2 = 1$. On peut donc écrire :

$$\begin{aligned}\nu_n(t)^2 &= \left(\sum_{j=1}^{D_{m'}} b_j \nu_n(\varphi_j)\right)^2, \\ &\leq \sum_j b_j^2 \sum_j \nu_n(\varphi_j)^2 = \sum_j \nu_n(\varphi_j)^2\end{aligned}$$

par l'inégalité de Cauchy-Schwarz, avec égalité pour le choix $b_j = \nu_n(\varphi_j)/\sqrt{\sum_j \nu_n^2(\varphi_j)}$. Ainsi,

$$\mathbb{E}\left[T_a\right] \leq \sum_{m' \in \mathcal{M}_n} \mathbb{E}\left[\left\{\sup_{t\in\mathcal{S}(m')} \nu_n^2(t) - \frac{V(m')}{6}\right\}_+\right].$$

Le Lemme 7 appliqué en remplaçant $m \vee m'$ par m', et $p(m, m')$ par $p(m', m')$ prouve que cette dernière quantité est bien majorée par C/n, pour le choix de $V(m') = 6 \times p(m', m')$, qui est celui indiqué à la ligne (5).

6.8.2. *Majoration de T_b^m.*

Pour majorer ce second terme, on distingue les cas $m' \leq m$ et $m' > m$:

$$\begin{aligned}
T_b^m &= \max_{m' \in \mathcal{M}_n} \left\{ \left\| h_{m\wedge m'} - \hat{h}^G_{m\wedge m'} \right\|^2 - \frac{V(m')}{6} \right\}_+ , \\
&= \max \left(\max_{\substack{m' \in \mathcal{M}_n \\ m' \leq m}} \left\{ \left\| h_{m'} - \hat{h}^G_{m'} \right\|^2 - \frac{V(m')}{6} \right\}_+ , \max_{\substack{m' \in \mathcal{M}_n \\ m' > m}} \left\{ \left\| h_m - \hat{h}^G_m \right\|^2 - \frac{V(m')}{6} \right\}_+ \right), \\
&\leq \max \left(T_a, \max_{\substack{m' \in \mathcal{M}_n \\ m' > m}} \left\{ \left\| h_m - \hat{h}^G_m \right\|^2 - \frac{V(m')}{6} \right\}_+ \right), \\
&\leq \max \left(T_a, \left\{ \left\| h_m - \hat{h}^G_m \right\|^2 - \frac{V(m)}{6} \right\}_+ \right), \\
&\leq T_a + \left\{ \left\| h_m - \hat{h}^G_m \right\|^2 - \frac{V(m)}{6} \right\}_+ ,
\end{aligned}$$

où l'on a utilisé que pour $m < m'$, $-V(m') \leq -V(m)$. Le paragraphe précédent prouve ensuite que T_a est majoré par C/n, et le Lemme 7 appliqué cette fois en substituant m à $m \vee m'$ entraîne a fortiori qu'il en est de même pour le second terme de la somme ci-dessus (puisque ce serait même vrai pour la somme sur $m \in \mathcal{M}_n$ de ce terme). Ainsi, $\mathbb{E}[T_b^m] \leq C/n$.

6.8.3. *Majoration de T_c^m.*

Ce terme n'est pas aléatoire, c'est un terme de biais. On le contrôle de la façon suivante :

$$\begin{aligned}
T_c^m &= \max_{m' \in \mathcal{M}_n} \left\| h_{m'} - h_{m\wedge m'} \right\|^2 , \\
&\leq \max_{\substack{m' \in \mathcal{M}_n \\ m \leq m'}} \left\| h_{m'} - h_m \right\|^2 , \\
&\leq 2 \max_{\substack{m' \in \mathcal{M}_n \\ m \leq m'}} \left\| h_{m'} - h \right\|^2 + 2 \left\| h - h_m \right\|^2 .
\end{aligned}$$

Or, si $m \leq m'$, $S_m \subset S_{m'}$, donc la projection orthogonale h_m de h sur S_m est aussi dans $S_{m'}$. Comme $h_{m'}$ est la projection de h sur $S_{m'}$, on a par conséquent,

$$\left\| h_{m'} - h \right\|^2 \leq \left\| h_m - h \right\|^2 .$$

Ainsi,

$$T_c^m \leq 4 \left\| h_m - h \right\|^2 .$$

□

6.9. **Estimation adaptative par pénalisation suite : une pénalité aléatoire.** Dans le Théorème 3 on a défini :

$$m \in \mathcal{M}_n \quad \mapsto \quad \mathrm{pen}_{\theta,\delta}(m) := 6\theta(1+2\delta)\phi_0^2 \mathbb{E}[Y_1^2] \frac{D_m}{n}.$$

Or l'espérance de Y_1^2 est inconnue. La question qui se pose est donc celle de son remplacement par un estimateur. On veut bien sûr conserver l'inégalité oracle pour l'estimateur sélectionné à partir d'une nouvelle pénalité entièrement connue. Le théorème suivant résout ce problème.

Théorème 4. *On se place sous les hypothèses du Théorème 3. Soient $\theta > 2$, $\delta > 0$, $0 < b < 1$ trois paramètres fixés. On note $\overline{Y_n^2} = (1/n)\sum_{i=1}^n Y_i^2$, et on définit l'application $\widehat{pen}_{\theta,\delta}$:*

$$m \in \mathcal{M}_n \quad \mapsto \quad \widehat{pen}_{\theta,\delta,b}(m) := \frac{6}{1-b}\theta(1+2\delta)\phi_0^2\overline{Y_n^2}\frac{D_m}{n}. \tag{13}$$

Soit aussi $\hat{m}^a \in argmin_{m\in\mathcal{M}_n}\gamma_n^G(\hat{h}_m^G) + \widehat{pen}_{\theta,\delta,b}(m)$.

Alors, l'estimateur $\hat{f}_{\hat{m}^a}^{G,G} = \hat{h}_{\hat{m}^a}^G \circ G$ vérifie l'inégalité-oracle suivante :

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G} - f\right\|_g^2\right] \quad \leq \quad \min_{m\in\mathcal{M}_n}\left\{\frac{\theta+2}{\theta-2}\left\|f - f_m^G\right\|_g^2 + \frac{2\theta(1+b)}{\theta-2}pen_{\theta,\delta}(m)\right\} + \frac{\theta^2}{\theta-2}\frac{C}{n}, \tag{14}$$

où l'on rappelle que $f_m^G = h_m \circ G$, h_m étant la projection orthogonale de h sur S_m, et où C est une constante ne dépendant que de ϕ_0^2, $\|f\|_\infty$, σ^2, $\mathbb{E}[|\varepsilon_1|^p]$, δ, $\mathbb{E}[Y_1^4]$, et $\mathbb{E}[Y_1^2]$.

Preuve

On notera pour simplifier $\widehat{\text{pen}}$ et pen les deux pénalités, en omettant de mettre en indice les paramètres dont elles dépendent.

On définit l'ensemble

$$\Omega_b = \left\{\left|\frac{\overline{Y_n^2}}{\mathbb{E}[Y_1^2]} - 1\right| < b\right\}.$$

On décompose ainsi le risque de l'estimateur :

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G} - f\right\|_g^2\right] = \mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G} - f\right\|_g^2 \mathbf{1}_{\Omega_b}\right] + \mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G} - f\right\|_g^2 \mathbf{1}_{\Omega_b^c}\right].$$

On a donc deux termes à majorer. L'idée est la suivante : sur Ω_b, l'estimateur $\overline{Y_n^2}$ est proche de la quantité qu'il estime, $\mathbb{E}[Y_1^2]$, donc la majoration du premier terme se ramène à la preuve du Théorème 3. Pour le second terme, l'idée est d'utiliser que la probabilité de Ω_b^c est petite, et donc de majorer $\|\hat{f}_{\hat{m}^a}^{G,G} - f\|_g^2$ par une quantité non aléatoire.

Majoration de $\mathbb{E}[\|\hat{f}_{\hat{m}^a}^{G,G} - f\|_g^2\mathbf{1}_{\Omega_b}]$

Notons déjà que sur Ω_b, on a les deux inégalités suivantes :

$$\mathbb{E}\left[Y_1^2\right] \leq \frac{1}{1-b}\overline{Y_n^2}, \quad \overline{Y_n^2} \leq (1+b)\mathbb{E}\left[Y_1^2\right].$$

On procède tout d'abord comme dans la preuve du théorème où la pénalité est déterministe : on obtient,

$$\begin{aligned}\left\|\hat{f}_{\hat{m}^a}^{G,G} - f\right\|_g^2 \quad \leq \quad & \frac{\theta}{\theta-2}\left\{\frac{\theta+2}{\theta}\left\|f - f_m^G\right\|_g^2 + \widehat{\text{pen}}(m) + \right. \\ & \left. \frac{\theta^2}{\theta-2}\sum_{m'\in\mathcal{M}_n}\left(\sup_{\substack{t\in S_{m\vee m'}\\ \|t\|=1}}(\nu_n(t))^2 - p(m,m')\right)_+ + \theta p(m,\hat{m}^a) - \widehat{\text{pen}}(\hat{m}^a)\right\}.\end{aligned}$$

où l'on rappelle les notations suivantes

$$p(m,\hat{m}^a) = 6(1+2\delta)\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_{m\vee\hat{m}^a}}{n},$$

$$\nu_n(t) = \frac{1}{n}\sum_{i=1}^n Y_i t\circ G(X_i) - \langle t\circ G, f\rangle_g.$$

Puis l'application de l'inégalité de Talagrand justifie l'existence d'une constante C telle que

$$\mathbb{E}\left[\sum_{m'\in\mathcal{M}_n}\left(\sup_{t\in S_{m\vee m'},\|t\|=1}(\nu_n(t))^2-p(m,m')\right)_+\right]\leq\frac{C}{n},$$

On obtient donc toujours

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2\mathbf{1}_{\Omega_b}\right]$$

$$\leq\frac{\theta+2}{\theta-2}\left\|f-f_m^G\right\|_g^2+\mathbb{E}\left[\mathbf{1}_{\Omega_b}\frac{\theta}{\theta-2}\widehat{\mathrm{pen}}(m)\right]+\frac{\theta^2}{\theta-2}\frac{C}{n}+\frac{\theta}{\theta-2}\mathbb{E}\left[\mathbf{1}_{\Omega_b}\left(\theta p(m,\hat{m}^b)-\widehat{\mathrm{pen}}\left(\hat{m}^a\right)\right)\right].$$

Mais l'on a, sur l'évènement Ω_b :

$$\begin{aligned}\theta p(m,\hat{m}^b) &= 6\theta(1+2\delta)\mathbb{E}\left[Y_1^2\right]\frac{D_m+D_{\hat{m}^a}}{n},\\ &\leq 6\theta(1+2\delta)\overline{Y_n^2}\frac{D_m+D_{\hat{m}^a}}{n}\\ &= \widehat{\mathrm{pen}}(m)+\widehat{\mathrm{pen}}(\hat{m}^a).\end{aligned}$$

On aboutit donc à

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2\mathbf{1}_{\Omega_b}\right]\leq\frac{\theta+2}{\theta-2}\left\|f-f_m^G\right\|_g^2+\frac{2\theta}{\theta-2}\mathbb{E}\left[\mathbf{1}_{\Omega_b}\widehat{\mathrm{pen}}(m)\right]+\frac{\theta^2}{\theta-2}\frac{C}{n}.$$

On utilise enfin, l'inégalité $\widehat{\mathrm{pen}}(m)\leq(1+b)\mathrm{pen}(m)$ valable sur Ω_b, pour conclure On aboutit donc à

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2\mathbf{1}_{\Omega_b}\right]\leq\frac{\theta+2}{\theta-2}\left\|f-f_m^G\right\|_g^2+\frac{2(1+b)\theta}{\theta-2}\mathrm{pen}(m)+\frac{\theta^2}{\theta-2}\frac{C}{n}.$$

Majoration de $\mathbb{E}[\|\hat{f}_{\hat{m}^a}^{G,G}-f\|_g^2\mathbf{1}_{\Omega_b^c}]$

On va montrer que ce terme est en $O(1/n)$, ce qui suffira. Commençons par majorer $\|\hat{f}_{\hat{m}^a}^{G,G}-f\|_g^2$. Partant de l'inégalité $\gamma_n^G(\hat{f}_{\hat{m}^a}^{G,G})\leq\gamma_n^G(f_{\hat{m}^a})$, on obtient classiquement

$$\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2\leq\|f_{\hat{m}^a}-f\|_g^2+2\nu_n\left(\hat{h}_{\hat{m}^a}^G-h_{\hat{m}^a}\right),$$

puis, pour un $\theta'>0$,

$$\begin{aligned}2\nu_n\left(\hat{h}_{\hat{m}^a}^G-h_{\hat{m}^a}\right) &\leq \frac{1}{\theta'}\left\|\hat{f}_{\hat{m}^a}^{G,G}-f_{\hat{m}^a}\right\|_g^2+\theta'\sup_{t\in S_{\hat{m}^a},\,\|t\|=1}\nu_n(t)^2,\\ &\leq \frac{2}{\theta'}\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2+\frac{2}{\theta'}\|f-f_{\hat{m}^a}\|_g^2+\theta'\sup_{t\in S_{\hat{m}^a},\,\|t\|=1}\nu_n(t)^2,\end{aligned}$$

D'où pour un $\theta'>2$,

$$\begin{aligned}\left\|\hat{f}_{\hat{m}^a}^{G,G}-f\right\|_g^2 &\leq \frac{\theta'+2}{\theta'-2}\|f_{\hat{m}^a}-f\|_g^2+\frac{\theta'^2}{\theta'-2}\sup_{t\in S_{\hat{m}^a},\,\|t\|=1}\nu_n(t)^2,\\ &\leq \frac{\theta'+2}{\theta'-2}\|f\|_g^2+\frac{\theta'^2}{\theta'-2}\sup_{t\in S_{\hat{m}^a},\,\|t\|=1}\nu_n(t)^2.\end{aligned}$$

On borne ensuite le supremum sur un ensemble aléatoire

$$\sup_{t\in S_{\hat{m}^a},\, \|t\|=1} \nu_n(t)^2 \leq \sum_{m\in\mathcal{M}_n}\left(\sup_{t\in S_m,\, \|t\|=1}\nu_n(t)^2 - \mathrm{pen}(m)\right)_+ + \mathrm{pen}(\hat{m}^a).$$

Par suite,

$$\begin{aligned}
&\mathbb{E}\left[\left\|\hat{f}^{G,G}_{\hat{m}^a} - f\right\|_g^2 \mathbf{1}_{\Omega_b^c}\right] \\
&\leq \frac{\theta'+2}{\theta'-2}\|f\|_g^2\,\mathbb{P}\left(\Omega_b^c\right) + \frac{\theta'^2}{\theta'-2}\left(\mathbb{E}\left[\sum_{m\in\mathcal{M}_n}\left(\sup_{t\in S_m,\, \|t\|=1}\nu_n(t)^2 - \mathrm{pen}(m)\right)_+\right] + \mathbb{E}\left[\mathrm{pen}(\hat{m}^a)\mathbf{1}_{\Omega_b^c}\right]\right).
\end{aligned}$$

On remarque que $\mathrm{pen}(\hat{m}^a) \leq 6\theta(1+2\delta)\phi_0^2\mathbb{E}[Y_1^2]$, en utilisant $D_{\hat{m}^a} \leq n$, ce qui entraîne,

$$\begin{aligned}
\mathbb{E}\left[\left\|\hat{f}^{G,G}_{\hat{m}^a} - f\right\|_g^2 \mathbf{1}_{\Omega_b^c}\right] &\leq \left(\frac{\theta'+2}{\theta'-2}\|f\|_g^2 + \frac{\theta'^2}{\theta'-2}6\theta(1+2\delta)\phi_0^2\mathbb{E}[Y_1^2]\right)\mathbb{P}\left(\Omega_b^c\right) \\
&\quad + \frac{\theta'^2}{\theta'-2}\mathbb{E}\left[\sum_{m\in\mathcal{M}_n}\left(\sup_{t\in S_m,\, \|t\|=1}\nu_n(t)^2 - \mathrm{pen}(m)\right)_+\right].
\end{aligned}$$

Il reste donc deux quantités à borner par un $O(1/n)$. Le Lemme 7 s'applique pour le contrôle du processus empirique, en remplaçant $\hat{m} \vee m$ par m, et en notant que $p(m,m) = \mathrm{pen}(m)$. On a donc

$$\mathbb{E}\left[\sum_{m\in\mathcal{M}_n}\left(\sup_{t\in S_m,\, \|t\|=1}\nu_n(t)^2 - \mathrm{pen}(m)\right)_+\right] = O\left(\frac{1}{n}\right).$$

Quant à la majoration de la probabilité du complémentaire de Ω_b, l'argument principal en est l'inégalité de Bienaymé-Tchebitcheff :

$$\begin{aligned}
\mathbb{P}\left(\Omega_b^c\right) &= \mathbb{P}\left(\left|\overline{Y_n^2} - \mathbb{E}\left[Y_1^2\right]\right| < b\mathbb{E}\left[Y_1^2\right]\right), \\
&\leq \frac{1}{b^2\mathbb{E}\left[Y_1^2\right]^2}\mathrm{Var}\left(\overline{Y_n^2}\right), \\
&\leq \frac{\mathbb{E}\left[Y_1^4\right]}{b^2\mathbb{E}\left[Y_1^2\right]^2}\frac{1}{n}.
\end{aligned}$$

Le dernier majorant étant en $O(1/n)$, cela termine la preuve.

□

Troisième partie 3. Estimation quand la loi du design est inconnue

7. Rappels

7.1. **Observations.** On rappelle que les observations sont constituées des couples i.i.d. $(X_i, Y_i)_{i=1,\ldots,n}$, avec $Y_i = f(X_i) + \varepsilon_i$ (les variables ε_i étant i.i.d. centrées de variance σ^2), ainsi que des variables X_{-i}, $i = 1, \ldots, n$, indépendantes entre elles, de même loi que les variables X_i (densité g, fonction de répartition G), et indépendantes des couples précédents. L'objectif est d'estimer la fonction de régression f, en commençant par l'estimation de $h = f \circ G^{-1}$.

7.2. **Modèles.** On considère les modèle trigonométriques, pour $m \in \mathcal{M}_n$,

$$S_m := \text{Vect}\left\{\varphi_1, \varphi_{2j}, \varphi_{2j+1}\ j = 1, \ldots, m\right\},$$

où, pour $x \in [0;1]$,

$$\begin{cases} \varphi_1(x) = 1, \\ \varphi_{2j}(x) = \sqrt{2}\cos(2\pi jx), \\ \varphi_{2j+1}(x) = \sqrt{2}\sin(2\pi jx). \end{cases}$$

On note toujours $D_m = 2m+1$ la dimension de S_m, et $m_{\max}$ l'indice du plus gros modèle de la collection. On rappelle qu'un tel modèle S_m vérifie l'hypothèse de connexion de normes avec $\phi_0 = 1$. De plus, les fonctions de la base étant de classe $\mathcal{C}^\infty$, on pourra appliquer la formule de Taylor-Lagrange aux ordres qui nous seront nécessaires. On utilisera aussi des majorations uniformes des dérivées successives :

$$\forall r \in \mathbb{N}\backslash\{0\},\ \forall j \in \{1, \ldots, n\},\ \|\varphi_j^{(r)}\|_\infty \leq \sqrt{2}(2\pi j)^r = \|\varphi_2^{(r)}\|_\infty j^r \leq \|\varphi_2^{(r)}\|_\infty D_m^r.$$

On remarque aussi que l'on a les égalités suivantes, pour $j \geq 1$:

$$\varphi_{2j}' = -2\pi j\varphi_{2j+1},\ \varphi_{2j}'' = -(2\pi j)^2\varphi_{2j},\ \varphi_{2j+1}' = 2\pi j\varphi_{2j+1},\ \varphi_{2j}'' = -(2\pi j)^2\varphi_{2j+1}.$$

7.3. **Espaces de régularité.** Dans ce cadre, on va faire l'hypothèse que la fonction h est dans un espace de Sobolev périodisé défini par

$$W^{per}(\beta, L) := \left\{h \in W_2^\beta(L),\ \forall j = 0, 1, \ldots, \beta - 1,\ h^{(j)}(0) = h^{(j)}(1)\right\}.$$

où l'on a,

$$W_2^\beta(L) := \left\{h : [0;1] \to \mathbb{R},\ h^{(\beta-1)} \text{ est absolument continue et } \int_0^1 \left(h^{(\beta)}(x)\right)^2 dx \leq L^2\right\},$$

le choix de tels espaces se justifiant par les propositions suivantes, qui nous seront utiles :

Proposition 9. *Soient $\beta \in \mathbb{N}\backslash\{0\}$, et $L > 0$. Une fonction $h = \sum_{j=1}^\infty \theta_j\varphi_j$ est dans l'espace $W^{per}(\beta, L)$ si et seulement si $(\theta_j)_j$ est dans l'ellipsoïde $\Theta(\beta, L^2/\pi^{2\beta})$.*

Rappelons que pour $\beta \in \mathbb{N}\backslash\{0\}$, et $Q > 0$, on note

$$\Theta(\beta, Q) := \left\{\theta \in \ell^2(\mathbb{N}),\ \sum_{j=1}^\infty \alpha_j^2\theta_j^2 \leq Q\right\},$$

où $(\alpha_j)_{j\in\mathbb{N}\backslash\{0\}}$ est définie par

$$\alpha_j = \begin{cases} j^\beta, \text{ si } j \text{ est pair,} \\ (j-1)^\beta, \text{ sinon.} \end{cases}$$

Remarquons que pour $0 \leq \beta' \leq \beta$, on a $\Theta(\beta, L^2/\pi^{2\beta}) \subset \Theta(\beta', L^2/\pi^{2\beta})$, et, par la proposition 9, l'inclusion analogue est valable pour les espaces de Sobolev périodisés : $W^{per}(\beta, L) \subset W^{per}(\beta', L)$, ce qui nous sera utile. De plus, cette proposition permet d'étendre la définition des $W^{per}(\beta, L)$ à des valeurs non entières de β : une fonction est dans un tel espace, si et seulement si la suite de ses coefficients de Fourier est dans l'ellipsoïde correspondant.

Une autre propriété justifie aussi l'usage de ces espaces :

Lemme 10. *Soient $L > 0$ et $h \in W^{per}(1, L) \cap \mathcal{C}^1$. On note Π_{S_m} l'opérateur de projection orthogonale sur l'espace S_m. Alors,*

$$\Pi_{S_m}(h') = (\Pi_{S_m}(h))'.$$

Enfin, une propriété d'approximation similaire au lemme 12 de Barron, Birgé et Massart est obtenue pour ces espaces : on en déduira des vitesses de convergences vers 0 du risque des estimateurs.

7.4. **Rappels sur la fonction de répartition empirique.** On note dans toute la suite, et pour tout $i \in \{1, \ldots, n\}$, $U_i = G(X_{-i})$, et $\hat{U}_n$ la fonction de répartition empirique associée au $n-$échantillon $(U_i)_{1\leq i\leq n}$. On rappelle le résultat suivant concernant la loi des U_i :

Lemme 11. *Quel que soit $i \in \{1, \ldots, n\}$, U_i suit la loi uniforme sur* $[0; 1]$.

Lemme 12. *Quel que soit $u \in \mathbb{R}$, on a l'égalité $\hat{G}_n(G^{-1}(u)) = \hat{U}_n(u)$.*

On utilisera l'inégalité suivante, qui est une version "uniforme" de l'inégalité de Hoeffding, c'est à dire qui permet de contrôler uniformément les déviations de la répartition empirique autour de son espérance.

Proposition 13. *(Inégalité de Dvoretzky-Kiefer-Wolfowitz) Il existe une constante $C > 0$, telle que, pour tout entier $n \geq 1$ et pour tout $\lambda > 0$,*

$$\mathbb{P}\left(\sup_{u\in\mathbb{R}}\left|\hat{U}_n(u) - u\right| \geq \lambda\right) \leq C\exp\left(-2n\lambda^2\right).$$

C'est précisément une version intégrée de cette inégalité qui va nous servir. On note désormais $\|\hat{U}_n - id\|_\infty$ à la place de $\sup_{u\in\mathbb{R}}\left|\hat{U}_n(u) - u\right|$.

Corollaire 14. *Pour tout $p \in \mathbb{N}\backslash\{0\}$, il existe une constante C_p telle que*

$$\mathbb{E}\left[\left\|\hat{U}_n - id\right\|_\infty^p\right] \leq \frac{C_p}{n^{p/2}}.$$

On utilisera aussi une autre version intégrée :

Corollaire 15. *Pour tout $\kappa > 0$, pour tout $p \in \mathbb{N}$, $p \geq 2$, il existe une constante numérique C, telle que,*

$$\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^p - \kappa\frac{\ln^{p/2}(n)}{n^{p/2}}\right)_+\right] \leq Cn^{-2^{\frac{2-p}{p}}\kappa^{2/p}}.$$

De plus, on a l'inégalité suivante :

$$\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa\frac{\ln(n)}{n}\right)^2\right] \leq Cn^{-2-2\kappa}.$$

Preuve du Corollaire 15

Pour la première inégalité, on note déjà que pour toute variable Z, $\mathbb{E}[Z_+] = \int_{\mathbb{R}_+}\mathbb{P}(Z_+ > t)dt$:

$$\begin{aligned}
&\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^p - \kappa\frac{\ln^{p/2}(n)}{n^{p/2}}\right)_+\right] \\
&\quad= \int_0^\infty \mathbb{P}\left(\left(\left\|\hat{U}_n - id\right\|_\infty^p - \kappa\frac{\ln^{p/2}(n)}{n^{p/2}}\right)_+ > t\right)dt, \\
&\quad\leq \int_0^\infty \mathbb{P}\left(\left\|\hat{U}_n - id\right\|_\infty^p - \kappa\frac{\ln^{p/2}(n)}{n^{p/2}} > t\right)dt, \\
&\quad= \int_0^\infty \mathbb{P}\left(\left\|\hat{U}_n - id\right\|_\infty > \left(t + \kappa\frac{\ln^{p/2}(n)}{n^{p/2}}\right)^{1/p}\right)dt,
\end{aligned}$$

On dispose de l'inégalité $(a+b)^{1/p} \geq 2^{1/p-1}(a^{1/p}+b^{1/p})$, valable pour $a,b>0$: en effet, on a par concavité de la fonction $u : x \mapsto x^{1/p}$,

$$\left(\frac{a+b}{2}\right)^{1/p} \geq \frac{1}{2}\left(a^{1/p}+b^{1/p}\right).$$

Par conséquent, en utilisant ceci, puis l'inégalité de D.K.W.,

$$\begin{aligned}
&\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^p - \kappa\frac{\ln^{p/2}(n)}{n^{p/2}}\right)_+\right] \\
&\quad \leq \int_0^\infty \mathbb{P}\left(\left\|\hat{U}_n - id\right\|_{\infty,[0;1]} > 2^{(1-p)/p}\left(t^{1/p} + \frac{\kappa^{1/p}\ln^{1/2}(n)}{n^{1/2}}\right)\right) dt \\
&\quad \leq C\int_0^\infty \exp\left\{-2n \times 2^{(2-2p)/p}\left(t^{1/p} + \frac{\kappa^{1/p}\ln^{1/2}(n)}{n^{1/2}}\right)^2\right\} dt, \\
&\quad \leq C\exp\left\{-2^{(2-p)/p}\kappa^{2/p}\ln(n)\right\}\int_0^\infty \exp\left\{-2^{(2-p)/p}nt^{2/p}\right\} dt, \qquad (15)\\
&\quad \leq Cn^{-2^{\frac{2-p}{p}}\kappa^{2/p}} \times 1,
\end{aligned}$$

où l'on a utilisé pour l'inégalité (15) que $(a+b)^2 \geq a^2+b^2$, pour $a,b>0$, et pour la dernière ligne de calcul, la majoration de l'intégrale restante par 1.
Pour la seconde inégalité, on note déjà que pour toute variable Z, $\mathbb{E}[Z^2] = \int_{\mathbb{R}_+} 2t\mathbb{P}(Z_+ > t)dt$, puis la majoration est du même type :

$$\begin{aligned}
&\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa\frac{\ln(n)}{n}\right)^2\right] \\
&\quad \leq \int_0^\infty 2t\mathbb{P}\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa\frac{\ln(n)}{n} > t\right) dt, \\
&\quad = \int_0^\infty 2t\mathbb{P}\left(\left\|\hat{U}_n - id\right\|_\infty > \sqrt{\kappa\frac{\ln(n)}{n} + t}\right) dt, \\
&\quad \leq C\int_0^\infty 2t\exp\left(-2n\left(\kappa\frac{\ln(n)}{n} + t\right)\right) dt, \\
&\quad = (C/2)n^{-2-2\kappa}\int_0^\infty u\exp(-u)du = (C/2)n^{-2-2\kappa}.
\end{aligned}$$

□

8. Estimation non adaptative

On fixe un indice $m \in \mathcal{M}_n$, et on estime la projection orthogonale de la fonction h sur le modèle S_m.

8.1. Construction de l'estimateur.

Dans le cas où G était connue, on considérait le contraste

$$\forall t \in L^2([a;b]),\ \gamma_n^G(t) := \|t\|_g^2 - \frac{2}{n}\sum_{i=1}^n Y_i t \circ G(X_i),$$

puis un estimateur minimisant ce contraste, $\hat{h}_m^G = \mathrm{argmin}_{t\in S_m}\gamma_n^G(t)$. On obtenait alors l'écriture explicite suivante :

$$\hat{h}_m^G = \sum_{j=1}^{D_m} \hat{a}_j^G \varphi_j, \text{ où } \forall j \in \{1,\dots,D_m\},\ \hat{a}_j^G = \frac{1}{n}\sum_{i=1}^n Y_i \varphi_j(G(X_i)),$$

et le calcul de l'espérance des $\hat{a}_j^G$ nous avait incité à poser $\hat{f}_m^{G,G} = \hat{h}_m^G \circ G$. Supposant maintenant que G est inconnue, on l'estime à l'aide de la répartition empirique calculée à partir de la moitié des observations :

$$\hat{G}_n :\ x \ \mapsto \ \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{X_{-i}\le x}.$$

On se sert de la technique de plug-in : on remplace G partout où elle intervient par son estimateur. On définit ainsi,

$$\forall j \in \{1,\dots,D_m\},\ \hat{a}_j^{\hat{G}} = \frac{1}{n}\sum_{i=1}^n Y_i\varphi_j(\hat{G}_n(X_i)),$$

puis

$$\hat{h}_m^{\hat{G}} = \sum_{j=1}^{D_m} \hat{a}_j^{\hat{G}}\varphi_j,$$

et enfin l'estimateur que l'on va étudier est

$$\hat{f}_m^{\hat{G},\hat{G}} = \hat{h}_m^{\hat{G}} \circ \hat{G}_n. \tag{16}$$

Remarquons que $\hat{h}_m^{\hat{G}}$ est aussi l'estimateur minimisant, sur l'espace S_m, le contraste $\gamma_n^{\hat{G}}$, défini pour $t \in L^2([0;1])$ par

$$\gamma_n^{\hat{G}}(t) = \|t\|^2 - \frac{2}{n}\sum_{i=1}^{D_m} Y_i t\circ \hat{G}_n(X_i). \tag{17}$$

Notre objectif est ici de montrer que sous certaines contraintes sur la valeur de D_m, et sur le choix du modèle S_m, la vitesse obtenue dans le cas où G était connue est conservée pour l'estimateur $\hat{f}_m^{\hat{G},\hat{G}}$.

8.2. **Résultats préliminaires à l'étude du risque.**

On rappelle les notations, pour $j \in \{1,\dots,D_m\}$,

$$a_j = \langle h, \varphi_j\rangle = \int_0^1 f\circ G^{-1}(u)\varphi_j(u)du, \quad \hat{a}_j^G = \frac{1}{n}\sum_{i=1}^n Y_i\varphi_j\circ G(X_i), \quad \hat{a}_j^{\hat{G}} = \frac{1}{n}\sum_{i=1}^n Y_i\varphi_j \circ \hat{G}_n(X_i).$$

Nous avons montré au chapitre précédent que $\mathbb{E}[\hat{a}_j^G] = a_j$, et nous utiliserons le calcul suivant :

Lemme 16. *Quel que soit* $j \in \{1,\dots,D_m\}$, $\mathbb{E}\left[\hat{a}_j^{\hat{G}} \mid (X'_l)_{l\in\{1,\dots,n\}}\right] = \int_0^1 f\circ G^{-1}(u)\varphi_j\circ \hat{U}_n(u)du.$

Preuve
Par linéarité de l'espérance conditionnelle, et en notant que les variables $Y_i\varphi_j \circ \hat{G}_n(X_i)$ sont

identiquement distribuées pour $i = 1, \dots, n$, conditionnellement à $\left(X'_l\right)_l$,

$$\begin{aligned}
\mathbb{E}\left[\hat{a}_j^{\hat{G}} \mid (X'_l)_l\right] &= \mathbb{E}\left[Y_1\varphi_j \circ \hat{G}_n(X_1) \mid (X'_l)_l\right], \\
&= \mathbb{E}\left[f(X_1)\varphi_j \circ \hat{G}_n(X_1) \mid (X'_l)_l\right] + \mathbb{E}\left[\varepsilon_1\varphi_j \circ \hat{G}_n(X_1) \mid (X'_l)_l\right], \\
&= \mathbb{E}\left[f(X_1)\varphi_j \circ \hat{G}_n(X_1) \mid (X'_l)_l\right] := \mathbb{E}\left[\chi(X_1, (X_{-l})_l) \mid (X'_l)_l\right],
\end{aligned}$$

où l'on a aussi utilisé que ε_1 et $\varphi_j \circ \hat{G}_n(X_1)$ sont indépendants conditionnellement à $(X_{-l})_l$, puis que l'espérance conditionnelle de ε_1 est nulle, car égale à son espérance (par indépendance avec (X'_l)). De plus, X_1 est indépendant de $(X_{-l})_l$, ce qui permet d'intégrer par rapport à sa loi pour calculer la dernière espérance :

$$\mathbb{E}\left[\hat{a}_j^{\hat{G}} \mid (X'_l)_{l\in\{1,\dots,n\}}\right] = \int_a^b \chi(x, (X_{-l})_l)g(x)dx = \int_0^1 f \circ G^{-1}(u)\varphi_j \circ \hat{G}_n(G^{-1})(u)du,$$

en posant $u = G(x)$. On conclut en utilisant le lemme 12 ci-dessus.

□

8.3. **Etude du risque.**

8.3.1. *Enoncé des résultats.* Pour évaluer le risque quadratique intégré de l'estimateur $\hat{f}_m^{\hat{G},\hat{G}}$, nous allons utiliser les résultats concernant le risque de l'estimateur construit dans le cas où la loi du design est supposée connue, $\hat{f}_m^{G,G}$, ainsi que l'intermédiaire suivant :

$$\hat{f}_m^{\hat{G},G} := \hat{h}_m^{\hat{G}} \circ G = \sum_{j=1}^{D_m} \hat{a}_j^{\hat{G}}\varphi_j \circ G.$$

Nous allons démontrer les résultats suivants :

Proposition 17. *Supposons que l'on travaille avec S_m le modèle trigonométrique classique, et que $h = f \circ G^{-1}$ est de classe $\mathcal{C}^1$ et dans $W^{per}(1, L)$ pour un certain $L > 0$. Alors, si $D_m \leq n^{1/3}/\ln(n)^{1/3}$, il existe une constante $C > 0$ (dépendant de L, $\|f\|_\infty$, $\|h\|$, $\|h'\|$, ϕ_0, $\|\varphi_2^{(r)}\|_\infty$, pour $r = 1, 2$ des constantes C_p, $p = 2, 3, 4, 6$ de l'inégalité de Dvoretzky-Kiefer-Wolfowitz et de $\mathbb{E}[Y_1^2]$) telles que*

$$\mathbb{E}\left[\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq 6\|f - f_m^G\|_g^2 + C\frac{D_m}{n}.$$

Proposition 18. *Supposons que l'on travaille avec S_m le modèle trigonométrique classique, et que $h = f \circ G^{-1}$ est de classe $\mathcal{C}^1$ et dans $W^{per}(1, L)$ pour un certain $L > 0$. Alors, si $D_m \leq n^{1/4}$, il existe une constante $C > 0$ (dépendant de L, $\|f\|_\infty$, $\|h\|$, $\|h'\|$, ϕ_0, $\|\varphi_2^{(r)}\|_\infty$, pour $r = 1, 2$ des constantes C_p, $p = 2, 3, 4, 6$ de l'inégalité de Dvoretzky-Kiefer-Wolfowitz et de $\mathbb{E}[Y_1^2]$) telles que*

$$\mathbb{E}\left[\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq 6\|f - f_m^G\|_g^2 + 36\phi_0^2\mathbb{E}[Y_1^2]\frac{D_m}{n} + C\frac{\ln(n)}{n}.$$

On en déduit un choix optimal pour D_m, dans le cas où la régularité (au sens des espaces de Sobolev) de la fonction h est connue. C'est l'objet des deux corollaires qui suivent.

Corollaire 19. *Supposons que l'on travaille avec S_m le modèle trigonométrique classique, et que $h = f \circ G^{-1}$ est de classe $\mathcal{C}^1$ et dans $W^{per}(\alpha, L)$ pour un certain $L > 0$, un certain $\alpha \geq 1$ et $D_m \leq n^{1/3}/\ln(n)^{1/3}$. Alors on a la majoration suivante, pour une certaine constante $C > 0$:*

$$\mathbb{E}\left[\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq 6\frac{L^2}{\pi^{2\alpha}}D_m^{-2\alpha} + C\frac{D_m}{n}.$$

L'hypothèse $\alpha \geq 1$ assure que $h \in W^{per}(1, L)$.

Corollaire 20. *Supposons que $h = f \circ G^{-1}$ est dans $W^{per}(\alpha, L)$ pour un certain $L > 0$, un certain $\alpha \geq 1$. Supposons aussi que h est de classe $\mathcal{C}^1$. Alors, le choix du modèle trigonométrique classique S_m, avec D_m d'ordre $n^{1/(2\alpha+1)}$, assure la majoration suivante du risque :*

$$\mathbb{E}\left[\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq C n^{-\frac{2\alpha}{2\alpha+1}},$$

où C est une constante.

Remarque : le choix $\alpha \geq 1$ assure que si D_m est de l'ordre $n^{1/(2\alpha+1)}$, alors $D_m \leq n^{1/3}/\ln(n)^{1/3}$, ce qui était exigé pour avoir la majoration du risque.

Introduction à la preuve de la proposition 17

On décompose la perte de notre estimateur de la façon suivante :

$$\begin{aligned}
\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g &\leq \|f - f_m^G\|_g + \|f_m^G - \hat{f}_m^{G,G}\|_g \\
&+ \left\|\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} - \mathbb{E}\left[\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} |(X_{-l})_l\right]\right\|_g \\
&+ \left\|\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} - \mathbb{E}\left[\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g \\
&+ \left\|\mathbb{E}\left[\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} |(X_{-l})_l\right]\right\|_g + \left\|\mathbb{E}\left[\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g .
\end{aligned}$$

On utilise ensuite la convexité de la fonction $x \mapsto x^2$, pour obtenir,

$$\left\|\hat{f}_m^{\hat{G},\hat{G}} - f\right\|_g^2 \leq 6 \sum_{l=0}^{4} T_l^m,$$

où

$$\begin{aligned}
T_0^m &= \|f - f_m^G\|_g^2 + \|f_m^G - \hat{f}_m^{G,G}\|_g^2, \\
T_1^m &= \left\|\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} - \mathbb{E}\left[\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} |(X_{-l})_l\right]\right\|_g^2, \\
T_2^m &= \left\|\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} - \mathbb{E}\left[\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g^2, \\
T_3^m &= \left\|\mathbb{E}\left[\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G} |(X_{-l})_l\right]\right\|_g^2, \\
T_4^m &= \left\|\mathbb{E}\left[\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g^2.
\end{aligned}$$

Le terme T_0^m est le majorant du risque dans le cas où G était supposée connue. Par conséquent, on l'a déjà traité :

$$\mathbb{E}\left[T_0^m\right] \leq C_\alpha D_m^{-2\alpha} + \phi_0^2 \mathbb{E}\left[Y_1^2\right] \frac{D_m}{n}.$$

Il s'agit donc de majorer d'une part $\mathbb{E}[T_l^m]$, $l = 1, 2$, et $\mathbb{E}[T_l^m]$, $l = 3, 4$ d'autre part. On utilise les lemmes suivants : dont les démonstrations font l'objet des prochains paragraphes :

Lemme 21. *Sous les hypothèses de la Proposition 17,*

$$\mathbb{E}\left[T_1^m\right] \leq C_2 \|\varphi_2'\|_\infty^2 \mathbb{E}\left[Y_1^2\right] \frac{D_m^3}{n^2}$$

Si $D_m \leq n^{1/2}$ en particulier,

$$\mathbb{E}\left[T_1^m\right] \leq C_2 \|\varphi_2'\|_\infty^2 \mathbb{E}\left[Y_1^2\right] \frac{D_m}{n}$$

Lemme 22. *Sous les hypothèses de la Proposition 17,*

$$\mathbb{E}\left[T_2^m\right] \leq C_2 \mathbb{E}\left[Y_1^2\right] \phi_0^2 \left\|\varphi_2'\right\|_\infty^2 \frac{D_m^4}{n^2}$$

Si $D_m \leq n^{1/3}$ en particulier,

$$\mathbb{E}\left[T_2^m\right] \leq C_2 \mathbb{E}\left[Y_1^2\right] \phi_0^2 \left\|\varphi_2'\right\|_\infty^2 \frac{D_m}{n}$$

Lemme 23. *Sous les hypothèses de la Proposition 17,*

$$\begin{aligned}\mathbb{E}[T_3^m] \quad \leq \quad & 6\left(C_2 \|h'\|^2 \frac{1}{n} + \phi_0^2 \mathbb{E}[Y_1^2] \frac{D_m}{n}\right) + 3C_4(\pi^4/4)\|h\|^2 \frac{D_m^4}{n^2} \\ & + \frac{C_6}{2} \left\|\varphi_2^{(3)}\right\|_\infty^2 \|h\|^2 \frac{D_m^7}{n^3}. \end{aligned} \tag{18}$$

En particulier, si $D_m \leq n^{1/3}$,

$$\mathbb{E}[T_3^m] \quad \leq \quad \left(6\left(C_2\|h'\|^2 + \phi_0^2 \mathbb{E}[Y_1^2]\right) + 3C_4(\pi^4/4)\|h\|^2 + \frac{C_6}{2}\left\|\varphi_2^{(3)}\right\|_\infty^2 \|h\|^2\right) \frac{D_m}{n}.$$

Lemme 24. *Sous les hypothèses de la Proposition 17,*

$$\mathbb{E}\left[T_4^m\right] \quad \leq \quad C\left(\frac{D_m^4 \ln(n)}{n^2} + \frac{D_m^7 \ln(n)}{n^3} + \frac{D_m^{10} \ln(n)}{n^4} + \frac{D_m^6}{n^4} + \frac{1}{n} + \frac{D_m^3}{n^2} + \frac{D_m^2}{n^{3/2}}\right)$$

où C est une constante fonction de $\|h\|, \|h'\|, \|\varphi_2'\|_{\infty,[0;1]}, \|\varphi_2^{(3)}\|_{\infty,[0;1]}, \|h'\|^2$, $\mathbb{E}[Y_1^2]$ et, ϕ_0^2. En particulier, si $D_m \leq n^{1/3}/\ln(n)^{1/3}$,

$$\mathbb{E}\left[T_4^m\right] \quad \leq \quad C\frac{D_m}{n}.$$

8.4. **Preuve du Lemme 21.**

Tout d'abord, la définition des estimateurs intervenants dans la définition de T_1^m permet d'écrire

$$T_1^m = \left\|\sum_{j=1}^{D_m}\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)\varphi_j \circ G\right\|_g^2.$$

En utilisant ensuite le lien entre $\|.\|_g$ et $\|.\|$, ainsi que l'orthonormalité des φ_j, on en déduit :

$$T_1^m = \sum_{j=1}^{D_m}\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2,$$

et donc,

$$\mathbb{E}\left[T_1^m\,|(X_{-l})_l\right] = \sum_{j=1}^{D_m} \mathrm{Var}\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right).$$

On calcule et majore chacune des variances de la somme. Pour $j \in \{1, \ldots, D_m\}$,

$$\mathrm{Var}\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right) \quad = \quad \frac{1}{n}\mathrm{Var}\left(Y_1\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)|(X_{-l})_l\right),$$

où l'on a utilisé que $\hat{a}_j^G - \hat{a}_j^{\hat{G}}$ s'écrit comme une moyenne des $Y_i(\varphi_j \circ G(X_i) - \varphi_j \circ \hat{G}_n(X_i))$ qui sont i.i.d. conditionnellement à $(X_{-l})_l$. On obtient ensuite

$$\begin{aligned}
\mathrm{Var}\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right) &\leq \frac{1}{n}\mathbb{E}\left[Y_1^2\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 |(X_{-l})_l\right],\\
&= \frac{1}{n}\mathbb{E}\left[f(X_1)^2\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 |(X_{-l})_l\right]\\
&\quad + \frac{1}{n}\mathbb{E}\left[\varepsilon_1^2\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 |(X_{-l})_l\right],\\
&= \frac{1}{n}\mathbb{E}\left[f(X_1)^2\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 |(X_{-l})_l\right]\\
&\quad + \frac{\sigma^2}{n}\mathbb{E}\left[\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 |(X_{-l})_l\right],
\end{aligned}$$

en utilisant les propriétés de ε_1 : successivement son indépendance avec toute fonction de X_1 et $(X_{-l})_l$, le fait qu'il est centré, et le fait que sa variance est notée σ^2. On utilise l'inégalité des accroissements finis pour écrire,

$$\left(\varphi_j \circ G(X_1) - \varphi_j \circ \hat{G}_n(X_1)\right)^2 \leq \|\varphi_j'\|_\infty^2 \left|G(X_1) - \hat{G}_n(X_1)\right|^2 \leq \|\varphi_j'\|_\infty^2 \left\|G - \hat{G}_n\right\|_\infty^2,$$

ce qui entraîne

$$\begin{aligned}
\mathbb{E}\left[T_1^m\,|(X_{-l})_l\right] &\leq \frac{1}{n}\left(\mathbb{E}\left[f^2(X_1)\right] + \sigma^2\right)\sum_{j=1}^{D_m}\|\varphi_j'\|_\infty^2 \left\|G - \hat{G}_n\right\|_\infty^2,\\
&= \frac{1}{n}\mathbb{E}\left[Y_1^2\right]\sum_{j=1}^{D_m}\|\varphi_j'\|_\infty^2 \left\|id - \hat{U}_n\right\|_\infty^2,
\end{aligned}$$

On majore la somme par $D_m \times (D_m\|\varphi_2'\|_\infty)^2$, pour obtenir,

$$\mathbb{E}\left[T_1^m\,|(X_{-l})_l\right] \leq \frac{D_m^3}{n}\|\varphi_2'\|_\infty^2 \mathbb{E}\left[Y_1^2\right]\left\|id - \hat{U}_n\right\|_\infty^2.$$

L'application du Corollaire 14 avec $p = 2$ permet de conclure

$$\mathbb{E}[T_1^m] \leq C_2\|\varphi_2'\|_\infty^2 \mathbb{E}\left[Y_1^2\right]\frac{D_m^3}{n^2}.$$

□

8.5. **Preuve du Lemme 22.**

On calcule

$$
\begin{aligned}
T_2^m &= \int_{[a;b]} \left(\hat{h}_m^{\hat{G}}(G(x)) - \hat{h}_m^{\hat{G}}\left(\hat{G}_n(x)\right) - \mathbb{E}\left[\hat{h}_m^{\hat{G}}(G(x)) - \hat{h}_m^{\hat{G}}\left(\hat{G}_n(x)\right) | (X_{-l})_l \right] \right)^2 g(x)dx, \\
&= \int_{[a;b]} \left\{ \sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right) \left(\varphi_j \circ G(x) - \varphi_j \circ \hat{G}_n(x) \right) \right\}^2 g(x)dx, \\
&= \int_{[0;1]} \left\{ \sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right) \left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u) \right) \right\}^2 du,
\end{aligned}
$$

Il vient ensuite, par l'inégalité de Cauchy-Schwarz, puis celle des accroissements finis,

$$
\begin{aligned}
T_2^m &\leq \int_{[0;1]} \sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right)^2 \sum_{j=1}^{D_m} \left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u) \right)^2 du, \\
&\leq \sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right)^2 \left\| \hat{U}_n - id \right\|_\infty^2 \sum_{j=1}^{D_m} \left\| \varphi_j' \right\|_\infty^2, \\
&\leq \left\| \varphi_2' \right\|_\infty^2 D_m^3 \left\| \hat{U}_n - id \right\|_\infty^2 \sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right)^2. \qquad (19)
\end{aligned}
$$

Par suite,

$$
\mathbb{E}\left[T_2^m\right] \leq \left\| \varphi_2' \right\|_\infty^2 D_m^3 \mathbb{E}\left[\left\| \hat{U}_n - id \right\|_\infty^2 \mathbb{E}\left[\sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right)^2 | (X_{-l})_l \right] \right].
$$

Or, $\mathbb{E}\left[\sum_{j=1}^{D_m} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} | (X_{-l})_l \right] \right)^2 | (X_{-l})_l \right]$

$$
\begin{aligned}
&= \sum_{j=1}^{D_m} \operatorname{Var}\left(\hat{a}_j^{\hat{G}} | (X_{-l})_l \right), \\
&= \frac{1}{n} \sum_{j=1}^{D_m} \operatorname{Var}\left(\varphi_j\left(\hat{G}_n(X_1) \right) Y_1 | (X_{-l})_l \right), \\
&\leq \frac{1}{n} \sum_{j=1}^{D_m} \mathbb{E}\left[\varphi_j^2\left(\hat{G}_n(X_1) \right) Y_1^2 | (X_{-l})_l \right], \\
&\leq \frac{1}{n} \left\| \sum_{j=1}^{D_m} \varphi_j \right\|_\infty^2 \mathbb{E}\left[Y_1^2 | (X_{-l})_l \right], \\
&\leq \frac{1}{n} \left\| \sum_{j=1}^{D_m} \varphi_j \right\|_\infty^2 \mathbb{E}\left[Y_1^2 \right], \\
&\leq \phi_0^2 \mathbb{E}\left[Y_1^2 \right] \frac{D_m}{n}.
\end{aligned}
$$

Ainsi, il reste à appliquer le Corollaire 14 :

$$\begin{aligned}\mathbb{E}\left[T_2^m\right] &\leq \mathbb{E}\left[Y_1^2\right]\phi_0^2\left\|\varphi_2'\right\|_\infty^2\frac{D_m^4}{n}\mathbb{E}\left[\left\|\hat{U}_n - id\right\|_\infty^2\right],\\ &= C_2\mathbb{E}\left[Y_1^2\right]\phi_0^2\left\|\varphi_2'\right\|_\infty^2\frac{D_m^4}{n^2}.\end{aligned}$$

□

8.6. **Preuve du Lemme 23.**

On a d'abord,

$$\begin{aligned}T_3^m &= \left\|\mathbb{E}\left[\hat{f}_m^{G,G} - \hat{f}_m^{\hat{G},G}\,|(X_l')_l\right]\right\|_g^2,\\ &= \sum_{j=1}^{D_m}\mathbb{E}\left[\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_l')_l\right]^2, \text{ en utilisant l'orthonormalité des } \varphi_j, \qquad (20)\end{aligned}$$

et donc,

$$T_3^m = \sum_{j=1}^{D_m}\left\{\int_0^1 f\circ G^{-1}(u)\left(\varphi_j(u) - \varphi_j\circ\hat{U}_n(u)\right)du\right\}^2,$$

en utilisant les calculs préliminaires d'espérance ci-dessus.
On effectue alors un développement de Taylor-Lagrange à l'ordre 3 pour chaque fonction φ_j ($j \in \{1,\ldots,D_m\}$). Il existe un réel aléatoire et dépendant de j, $\hat{\alpha}_{j,n,u}$, tel que

$$\begin{aligned}&\varphi_j(u) - \varphi_j(\hat{U}_n(u))\\ &\quad= \varphi_j'(u)(u-\hat{U}_n(u)) - \varphi_j''(u)\frac{(u-\hat{U}_n(u))^2}{2} + \varphi_j^{(3)}(\hat{\alpha}_{j,n,u})\frac{(u-\hat{U}_n(u))^3}{6}. \qquad (21)\end{aligned}$$

On en déduit une majoration en trois termes :

$$T_3^m \leq 3T_{3,1}^m + 3T_{3,2}^m + 3T_{3,3}^m,$$

où l'on a noté

$$\begin{aligned}T_{3,1}^m &= \sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u) - u\right)\varphi_j'(u)du\right\}^2,\\ T_{3,2}^m &= (1/4)\sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u) - u\right)^2\varphi_j''(u)du\right\}^2,\\ T_{3,3}^m &= (1/6)\sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u) - u\right)^3\varphi_j^{(3)}(\hat{\alpha}_{j,n,u})du\right\}^2.\end{aligned}$$

Pour majorer le premier terme, l'idée est d'abord d'écrire,

$$u - \hat{U}_n(u) = \frac{-1}{n}\sum_{i=1}^n \mathbf{1}_{U_i\leq u} + \mathbb{E}\left[\mathbf{1}_{U_i\leq u}\right],$$

ce qui entraîne :

$$T_{3,1}^m = \sum_{j=1}^{D_m} \left(\frac{1}{n} \sum_{i=1}^{n} A_{i,j} - \mathbb{E}[A_{i,j}] \right)^2, \quad \text{où } A_{i,j} = \int_{U_i}^{1} h(u)\varphi_j'(u)du.$$

On intègre alors par parties pour calculer $A_{i,j}$ (on prend une primitive de φ_j', et on dérive $h = f \circ G^{-1}$, que l'on a supposé $\mathcal{C}^1$), et l'on obtient, quel que soit $j \in \{1, \ldots, D_m\}$, et $i \in \{1, \ldots, n\}$,

$$A_{i,j} - \mathbb{E}[A_{i,j}] = -h(U_i)\varphi_j(U_i) + \mathbb{E}\left[h(U_i)\varphi_j(U_i)\right] - \int_0^1 \left(\mathbf{1}_{U_i \le u} - \mathbb{E}\left[\mathbf{1}_{U_i \le u}\right]\right) h'(u)\varphi_j(u)du.$$

Par suite, en sommant sur i, élevant au carré, et sommant sur j, on dispose à nouveau une majoration en deux termes,

$$T_{3,1}^m \le 2T_{3,1,1}^m + 2T_{3,1,2}^m,$$

avec les notations

$$(22) \qquad \begin{aligned} T_{3,1,1}^m &= \sum_{j=1}^{D_m} \left\{ \frac{1}{n} \sum_{i=1}^{n} h(U_i)\varphi_j(U_i) - \mathbb{E}\left[h(U_i)\varphi_j(U_i)\right] \right\}^2, \\ T_{3,1,2}^m &= \sum_{j=1}^{D_m} \left\{ \int_0^1 h'(u) \left(\hat{U}_n(u) - u \right) \varphi_j(u) du \right\}^2. \end{aligned}$$

L'espérance du terme $T_{3,1,1}^m$ est une somme sur j de variances de moyennes empiriques, que l'on calcule et majore comme pour T_1^m et T_2^m :

$$\begin{aligned} \mathbb{E}\left[T_{3,1,1}^m\right] &\le \frac{1}{n} \sum_{j=1}^{D_m} \mathbb{E}\left[(h(U_1)\varphi_j(U_1))^2\right] \le \frac{1}{n} \left\| \sum_{j=1}^{D_m} \varphi_j^2 \right\|_\infty \int_0^1 h(u)^2 du, \\ &= \int_0^1 h(u)^2 du \phi_0^2 \frac{D_m}{n} = \phi_0^2 \mathbb{E}[f(X_1)^2] \frac{D_m}{n} \le \phi_0^2 \mathbb{E}[Y_1^2] \frac{D_m}{n}, \end{aligned}$$

en utilisant la connexion de normes, et la définition de h.
Quant au terme $T_{3,1,2}^m$, on remarque qu'il s'agit de la norme d'une projection sur S_m, et on le majore en utilisant que la norme du projeté est inférieure ou égale à la norme de la fonction que l'on projette :

$$\begin{aligned} T_{3,1,2}^m &= \sum_{j=1}^{D_m} \left(\langle h'(\hat{U}_n - id), \varphi_j \rangle \right)^2 = \left\| \Pi_{S_m}(h'(\hat{U}_n - id)) \right\|^2, \\ &\le \left\| h'(\hat{U}_n - id) \right\|^2 \le \|h'\|^2 \|\hat{U}_n - id\|_\infty^2. \end{aligned}$$

En utilisant le Corollaire 14 pour conclure, on a $\mathbb{E}[T_{3,1,2}^m] \le C_2 \|h'\|^2/n$. Par suite,

$$\mathbb{E}\left[T_{3,1}^m\right] \le 2\left(C_2 \|h'\|^2 \frac{1}{n} + \phi_0^2 \mathbb{E}[Y_1^2] \frac{D_m}{n} \right).$$

Pour majorer le terme $T_{3,2}^m$, on commence par remarquer que pour $j \geq 2$, $\varphi_j'' = -(\pi\mu_j)^2\varphi_j$, où l'on note $\mu_j = j$ si j est pair, et $\mu_j = j-1$ si j est impair. Par conséquent,

$$\begin{aligned}
\mathbb{E}\left[T_{3,2}^m\right] &= (\pi^4/4)\mathbb{E}\left[\sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u)-u\right)^2\mu_j^2\varphi_j(u)du\right\}^2\right], \\
&\leq (\pi^4/4)D_m^4\mathbb{E}\left[\sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u)-u\right)^2\varphi_j(u)du\right\}^2\right], \\
&= (\pi^4/4)D_m^4\mathbb{E}\left[\sum_{j=1}^{D_m}\left\{\langle h\left(\hat{U}_n-id\right)^2,\varphi_j\rangle\right\}^2\right], \\
&= (\pi^4/4)D_m^4\mathbb{E}\left[\left\|\Pi_{S_{m_1}^{(1)}}h\left(\hat{U}_n-id\right)^2\right\|^2\right], \\
&\leq (\pi^4/4)D_m^4\mathbb{E}\left[\left\|h\left(\hat{U}_n-id\right)^2\right\|^2\right], \\
&\leq (\pi^4/4)D_m^4\mathbb{E}\left[\left\|\hat{U}_n-id\right\|_\infty^4\right]\int_{[0;1]}h^2(u)du, \\
&\leq C_4(\pi^4/4)\|h\|^2\frac{D_m^4}{n^2}.
\end{aligned}$$

Enfin, pour le dernier terme,

$$\begin{aligned}
\mathbb{E}\left[T_{3,3}^m\right] &= (1/6)\mathbb{E}\left[\sum_{j=1}^{D_m}\left\{\int_0^1 h(u)\left(\hat{U}_n(u)-u\right)^3\varphi_j^{(3)}(\hat{\alpha}_{j,n,u})du\right\}^2\right], \\
&\leq (1/6)\sum_{j=1}^{D_m}\left\|\varphi_j^{(3)}\right\|_\infty^2\|h\|^2\mathbb{E}\left[\left\|\hat{U}_n-id\right\|_\infty^6\right], \\
&\leq \frac{C_6}{6}\left\|\varphi_2^{(3)}\right\|_\infty^2\|h\|^2\frac{D_m^7}{n^3}.
\end{aligned}$$

En résumé,

$$\begin{aligned}
\mathbb{E}[T_3^m] \leq\ & 6\left(C_2\|h'\|^2\frac{1}{n}+\phi_0^2\mathbb{E}[Y_1^2]\frac{D_m}{n}\right)+3C_4(\pi^4/4)\|h\|^2\frac{D_m^4}{n^2} \\
&+\frac{C_6}{2}\left\|\varphi_2^{(3)}\right\|_\infty^2\|h\|^2\frac{D_m^7}{n^3}.
\end{aligned} \tag{23}$$

□

8.7. **Preuve du Lemme 24.**

On commence à décomposer le dernier terme de la façon suivante :

$$\begin{aligned}
T_4^m &= \left\|\mathbb{E}\left[\hat{f}_m^{\hat{G},G} - \hat{f}_m^{\hat{G},\hat{G}}\,|(X_{-l})_l\right]\right\|_g^2, \\
&= \left\|\mathbb{E}\left[\sum_{j=1}^{D_m} \hat{a}_j^{\hat{G}}(\varphi_j \circ G - \varphi_j \circ \hat{G}_n)\,|(X_{-l})_l\right]\right\|_g^2, \\
&\leq 2\left\|\mathbb{E}\left[\sum_{j=1}^{D_m}(\hat{a}_j^{\hat{G}} - a_j)(\varphi_j \circ G - \varphi_j \circ \hat{G}_n)\,|(X_{-l})_l\right]\right\|_g^2 + 2\left\|\mathbb{E}\left[\sum_{j=1}^{D_m} a_j(\varphi_j \circ G - \varphi_j \circ \hat{G}_n)\,|(X_{-l})_l\,|\right]\right\|_g^2, \\
&:= 2T_{4,1}^m + 2T_{4,2}^m.
\end{aligned}$$

Commençons par majorer $T_{4,1}^m$:

$$\begin{aligned}
T_{4,1}^m &\leq \int_{[a;b]} \mathbb{E}\left[\left(\sum_{j=1}^{D_m}\left(\hat{a}_j^{\hat{G}} - a_j\right)\left(\varphi_j \circ G(x) - \varphi_j \circ \hat{G}_n(x)\right)\right)^2 |(X_{-l})_l\right] g(x)dx, \\
&\leq \int_{[a;b]} \mathbb{E}\left[\sum_{j=1}^{D_m}\left(\hat{a}_j^{\hat{G}} - a_j\right)^2 \sum_{j=1}^{D_m}\left(\varphi_j \circ G(x) - \varphi_j \circ \hat{G}_n(x)\right)^2 |(X_{-l})_l\right] g(x)dx, \\
&= \mathbb{E}\left[\sum_{j=1}^{D_m}\left(\hat{a}_j^{\hat{G}} - a_j\right)^2 \sum_{j=1}^{D_m}\int_{[0;1]}\left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u)\right)^2 du\,|(X_{-l})_l\right], \\
&\leq 2T_{4,1,1}^m + 2T_{4,1,2}^m,
\end{aligned}$$

où l'on note

$$\begin{aligned}
T_{4,1,1}^m &= \int_{[0;1]} \mathbb{E}\left[\left\{\sum_{j=1}^{D_m}\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2\right\}\left\{\sum_{j=1}^{D_m}\left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u)\right)^2\right\}|(X_{-l})_l\right] du, \\
T_{4,1,2}^m &= \int_{[0;1]} \mathbb{E}\left[\left\{\sum_{j=1}^{D_m}\left(\mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right] - a_j\right)^2\right\}\left\{\sum_{j=1}^{D_m}\left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u)\right)^2\right\}|(X_{-l})_l\right] du
\end{aligned}$$

Le premier terme s'écrit aussi

$$T_{4,1,1}^m = \sum_{j=1}^{D_m} \mathrm{Var}\left(\hat{a}_j^{\hat{G}}|(X_{-l})_l\right)\int_{[0;1]}\sum_{j=1}^{D_m}\left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u)\right)^2 du,$$

et le second se simplifie en

$$T_{4,1,2}^m = T_3^{m,b}\int_{[0;1]}\sum_{j=1}^{D_m}\left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u)\right)^2 du$$

d'après la ligne de calcul (20), en notant

$$T_3^{m,b} = \sum_{j=1}^{D_m}\left(\mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right] - a_j\right)^2. \tag{24}$$

Puis on utilise l'inégalité des accroissements finis dans chacun des deux termes pour obtenir :

$$\begin{aligned}\sum_{j=1}^{D_m}\int_{[0;1]}\left(\varphi_j(u)-\varphi_j\circ\hat{U}_n(u)\right)^2du &\leq \sum_{j=1}^{D_m}\|\varphi_j\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2,\\ &\leq D_m^3\|\varphi_2'\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2.\end{aligned}$$

Cela nous donne donc,

$$\begin{aligned}&T_{4,1,1}^m\leq\sum_{j=1}^{D_m}\mathrm{Var}\left(\hat{a}_j^{\hat{G}}|(X_{-l})_l\right)D_m^3\|\varphi_2'\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2,\\ &T_{4,1,2}^m\leq T_3^{m,b}D_m^3\|\varphi_2'\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2.\end{aligned}$$

Pour terminer la majoration de $T_{4,1,1}^m$, on contrôle la variance conditionnelle :

$$\begin{aligned}\mathrm{Var}\left(\hat{a}_j^{\hat{G}}|\,(X_{-l})_l\right) &= \mathrm{Var}\left\{\frac{1}{n}\sum_{i=1}^{n}Y_i\varphi_j\circ\hat{G}_n(X_i)|\,(X_{-l})_l\right\},\\ &= \frac{1}{n}\mathrm{Var}\left\{Y_1\varphi_j\circ\hat{G}_n(X_1)|\,(X_{-l})_l\right\},\\ &\leq \frac{1}{n}\mathbb{E}\left[Y_1^2\left(\varphi_j\circ\hat{G}_n(X_1)\right)^2|\,(X_{-l})_l\right],\\ &\leq \frac{1}{n}\|\varphi_j\|_\infty^2\left(\mathbb{E}\left[f^2(X_1)\right]+\sigma^2\right)=\frac{1}{n}\|\varphi_j\|_\infty^2\,\mathbb{E}\left[Y_1^2\right].\end{aligned}$$

Donc la connexion de normes entraîne

$$\sum_{j=1}^{D_m}\mathrm{Var}\left(\hat{a}_j^{\hat{G}}|\,(X_{-l})_l\right)\leq\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m}{n},$$

et par suite,

$$T_{4,1,1}^m\leq\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m}{n}\times D_m^3\|\varphi_2'\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2,\tag{25}$$

ce qui donne, en intégrant et en appliquant l'inégalité de D.K.W :

$$\mathbb{E}\left[T_{4,1,1}^m\right]\leq\phi_0^2\mathbb{E}\left[Y_1^2\right]C_2\|\varphi_2'\|_{\infty,[0;1]}^2\frac{D_m^4}{n^2}.$$

Pour majorer $T_{4,1,2}^m$, on doit être plus précis sur la majoration de $\|\hat{U}_n-id\|_\infty$: on introduit dans cette dernière majoration la somme des indicatrices des évènements $\{\|\hat{U}_n-id\|_\infty\leq\alpha_n\}$, et $\{\|\hat{U}_n-id\|_\infty>\alpha_n\}$, où $\alpha_n=\sqrt{2\ln(n)/n}$, ce qui donne le majorant :

$$T_{4,1,2}^m\leq T_{4,1,2,1}^m+T_{4,1,2,2}^m,$$

avec

$$\begin{aligned}&T_{4,1,2,1}^m=D_m^3\|\varphi_2'\|_\infty^2\alpha_n^2T_3^{m,b},\\ &T_{4,1,2,2}^m=D_m^3\|\varphi_2'\|_\infty^2\left\|\hat{U}_n-id\right\|_\infty^2\mathbf{1}_{\|\hat{U}_n-id\|_\infty>\alpha_n}T_3^{m,b}.\end{aligned}$$

Pour $T^m_{4,1,2,1}$, on utilise la majoration de l'espérance de $T_3^{m,b}$ obtenue au paragraphe précédent :

$$\begin{aligned}
\mathbb{E}\left[T^m_{4,1,2,1}\right] &\leq D_m^3\|\varphi_2'\|_\infty^2\alpha_n^2 \times C\left(\frac{D_m}{n}+\frac{D_m^7}{n^3}+\frac{D_m^4}{n^2}\right),\\
&= C\left(\frac{D_m^4\ln(n)}{n^2}+\frac{D_m^{10}\ln(n)}{n^4}+\frac{D_m^7\ln(n)}{n^3}\right),
\end{aligned}$$

où la constante C dépend ici de $\|h\|$, $\|h'\|$, $\|\varphi_2'\|_\infty$, $\|\varphi_2^{(3)}\|_\infty$, $\mathbb{E}[Y_1^2]$, et ϕ_0^2. Enfin, pour $T^m_{4,1,2,2}$, on commence par majorer grossièrement $T_3^{m,b}$:

$$\begin{aligned}
\left(\mathbb{E}\left[\hat{a}_j^{\hat{G}} \,|\, (X_{-l})_l\right]-a_j\right)^2 &= \left(\int_0^1 h(u)\left(\varphi_j(u)-\varphi_j\circ\hat{U}_n(u)\right)du\right)^2,\\
&\leq \int_0^1 (h(u))^2\left(\varphi_j(u)-\varphi_j\circ\hat{U}_n(u)\right)^2 du,\\
&\leq \|\varphi_j'\|_\infty^2\int_0^1 h^2(u)\left|\hat{U}_n(u)-u\right|^2 du,\\
&\qquad \text{en utilisant l'inégalité des accroissements finis,}\\
&\leq \|h\|^2\|\varphi_j'\|_\infty^2\left\|id-\hat{U}_n\right\|_\infty^2.
\end{aligned}$$

Ce qui donne en sommant sur j,

$$T_3^{m,b}=\sum_{j=1}^{D_m}\left(\mathbb{E}\left[\hat{a}_j^{\hat{G}} \,|\, (X_{-l})_l\right]-a_j\right)^2 \leq \|h\|^2\left\|\varphi_2'\right\|_\infty^2\left\|id-\hat{U}_n\right\|_\infty^2 D_m^3.$$

Ainsi, par les inégalités de Cauchy-Schwarz, de la Proposition 13 et du Corollaire 14,

$$\begin{aligned}
\mathbb{E}\left[T^m_{4,1,2,2}\right] &\leq D_m^6\|h\|^2\|\varphi_2'\|_\infty^4\,\mathbb{E}\left[\left\|\hat{U}_n-id\right\|_\infty^4\mathbf{1}_{\|\hat{U}_n-id\|_\infty>\alpha_n}\right],\\
&\leq D_m^6\|h\|^2\|\varphi_2'\|_\infty^4\frac{\sqrt{C_8}}{n^2}\exp\left(-n\alpha_n^2\right),\\
&= \|h\|^2\sqrt{C_8}\|\varphi_2'\|_{\infty,[0;1]}^4\frac{D_m^6}{n^4}.
\end{aligned}$$

Continuons avec la majoration de $T^m_{4,2}$. On commence par écrire :

$$\begin{aligned}
T^m_{4,2} &= \left\|\mathbb{E}\left[\sum_{j=1}^{D_m} a_j(\varphi_j\circ G-\varphi_j\circ\hat{G}_n)\,|(X_{-l})_l\right]\right\|_g^2,\\
&= \left\|\sum_{j=1}^{D_m} a_j(\varphi_j\circ G-\varphi_j\circ\hat{G}_n)\right\|_g^2,
\end{aligned}$$

et donc

$$\mathbb{E}\left[T_{4,2}^m\right] \leq \mathbb{E}\left[\left\|\sum_{j=1}^{D_m} a_j(\varphi_j \circ G - \varphi_j \circ \hat{G}_n)\right\|_g^2\right], \tag{26}$$

$$= \mathbb{E}\left[\sum_{j,k=1}^{D_m} a_j a_k \int_0^1 (\varphi_j(u) - \varphi_j \circ \hat{U}_n(u))(\varphi_k(u) - \varphi_k \circ \hat{U}_n(u))du\right].$$

On effectue ensuite le développement de Taylor à l'ordre 2 pour chacun des deux facteurs dont le produit est l'intégrande ci-dessus. Ceci fournit la décomposition :

$$\mathbb{E}\left[T_{4,2}^m\right] \leq \mathbb{E}\left[T_{4,2,1}^m + T_{4,2,2}^m + T_{4,2,3}^m\right],$$

avec

$$T_{4,2,1}^m = \sum_{j,k=1}^{D_m} a_j a_k \int_0^1 (u - \hat{U}_n(u))^2 \varphi_j'(u)\varphi_k'(u)du,$$
$$T_{4,2,2}^m = (1/4) \sum_{j,k=1}^{D_m} a_j a_k \int_0^1 (u - \hat{U}_n(u))^4 \varphi_j''(\hat{\alpha}_{j,n,u})\varphi_k''(\hat{\alpha}_{k,n,u})du,$$
$$T_{4,2,3}^m = \sum_{j,k=1}^{D_m} a_j a_k \int_0^1 (u - \hat{U}_n(u))^3 \varphi_j''(\hat{\alpha}_{j,n,u})\varphi_k'(u)du,$$

et l'on majore à nouveau chacun des trois termes, en rappelant que $a_l = \langle h, \varphi_l\rangle$. Ainsi, le premier terme s'écrit aussi

$$T_{4,2,1}^m = \int_0^1 (u - \hat{U}_n(u))^2 \left\{(\Pi_{S_m}(h))'(u)\right\}^2 du,$$
$$= \int_0^1 (u - \hat{U}_n(u))^2 \left\{\Pi_{S_m}(h')(u)\right\}^2 du,$$
$$\leq \left\|id - \hat{U}_n\right\|_\infty^2 \left\|\Pi_{S_m}(h')\right\|^2 \leq \left\|id - \hat{U}_n\right\|_\infty^2 \left\|h'\right\|^2, \tag{27}$$

où l'on a utilisé le fait que la dérivée de la projection de h coïncide avec la projection de h'. Donc, par l'inégalité de D.K.W.,

$$\mathbb{E}\left[T_{4,2,1}^m\right] \leq \frac{C_2\|h'\|^2}{n}.$$

D'autre part, le second terme est

$$T_{4,2,2}^m = (1/4) \int_0^1 (u - \hat{U}_n(u))^4 \left(\sum_{j=1}^{D_m} a_j \varphi_j''(\hat{\alpha}_{j,n,u})\right)^2 du. \tag{28}$$

Cette fois, il n'apparait pas une dérivée de projection : en effet, le point d'évaluation des fonctions φ_j'' dépend lui même de j. On va exploiter le fait que $h \in W^{per}(1, L)$: on introduit les réels μ_j, définis pour $j \in \{1, \ldots, D_m\}$ par

$$\mu_j = \begin{cases} j & \text{si } j \text{ est pair,} \\ j-1 & \text{sinon.} \end{cases}$$

Ceci permet d'écrire,

$$\begin{aligned}\left(\sum_{j=1}^{D_m} a_j \varphi_j''(\hat{\alpha}_{j,n,u})\right)^2 &= \|\varphi_2''\|_\infty^2 \left(\sum_{j=1}^{D_m} a_j \mu_j^2\right)^2, \\ &\leq \|\varphi_2''\|_\infty^2 \left(\sum_{j=1}^{D_m} a_j^2 \mu_j^2\right) \sum_{j=1}^{D_m} \mu_j^2, \text{ par l'inégalité de Cauchy-Schwarz,} \quad (29) \\ &\leq \|\varphi_2''\|_\infty^2 \frac{L^2}{\pi^2} \times D_m^3\end{aligned}$$

où l'on utilise la proposition 9 (la suite $(a_j)_j$ est dans l'ellipsoïde $\Theta(1, L^2/\pi^2)$) et la majoration de μ_j par D_m. Ainsi,

$$T_{4,2,2}^m \leq (1/4)\|\varphi_2''\|_\infty^2 \frac{L^2}{\pi^2} \times D_m^3 \left\|\hat{U}_n - id\right\|_\infty^4.$$

Appliquant aussi l'inégalité de D.K.W. intégrée avec $p = 4$, on en déduit,

$$\mathbb{E}\left[T_{4,2,2}^m\right] \leq (1/4)\frac{C_4}{n^2}\|\varphi_2''\|_\infty^2 \frac{L^2}{\pi^2} \times D_m^3.$$

Enfin, pour le dernier terme :

$$T_{4,2,3}^m = \int_0^1 (u - \hat{U}_n(u))^3 \left(\sum_{j=1}^{D_m} a_j \varphi_j''(\hat{\alpha}_{j,n,u})\right)\left(\sum_{k=1}^{D_m} a_k \varphi_k'(u)\right) du,$$

le calcul que l'on vient d'effectuer prouve que

$$\left|\sum_{j=1}^{D_m} a_j \varphi_j''(\hat{\alpha}_{j,n,u})\right| \leq \|\varphi_2''\|_\infty \frac{L}{\pi} D_m^{3/2},$$

et la même technique de calcul entraîne, pour $u \in [0; 1]$,

$$\left|\sum_{k=1}^{D_m} a_k \varphi_k'(u)\right| \leq \|\varphi_2'\|_\infty \frac{L}{\pi} D_m^{1/2}.$$

Donc,

$$T_{4,2,3}^m \leq \|\varphi_2'\|_\infty \|\varphi_2''\|_\infty \frac{L^2}{\pi^2} D_m^2 \left\|\hat{U}_n - id\right\|_\infty^3.$$

Ainsi, toujours par l'inégalité de D.K.W. (avec $p = 3$),

$$\mathbb{E}\left[T_{4,2,3}^m\right] \leq \|\varphi_2'\|_\infty \|\varphi_2''\|_\infty \frac{L^2}{\pi^2} C_3 \frac{D_m^2}{n^{3/2}}.$$

Finalement, en notant C une constante, on a la majoration

$$\mathbb{E}\left[T_4^m\right] \leq C\left(\frac{D_m^4 \ln(n)}{n^2} + \frac{D_m^7 \ln(n)}{n^3} + \frac{D_m^{10} \ln(n)}{n^4} + \frac{D_m^6}{n^4} + \frac{1}{n} + \frac{D_m^3}{n^2} + \frac{D_m^2}{n^{3/2}}\right).$$

□

9. Estimation adaptative

On ne suppose plus fixé un indice $m \in \mathcal{M}_n$. Disposant de la collection $(\hat{f}_m^{\hat{G}})_{m\in\mathcal{M}_n}$ d'estimateurs, on cherche à sélectionner un indice m convenable sur la base des observations.
On propose comme dans le cas où la loi du design était supposée connue, deux méthodes, l'une fondée sur un procédure classique de sélection de modèles par pénalisation, et l'autre fondée sur la comparaison des estimateurs ci-dessus deux à deux, basée sur des travaux de Lepski et Goldenshluger.

9.1. **Construction de l'estimateur par méthode de Lepski.**

On définit les deux quantités suivantes, quel que soit $m \in \mathcal{M}_n$:

$$\begin{aligned} &V^{bis}(m) = c\phi_0^2\mathbb{E}\left[Y_1^2\right]\tfrac{D_m}{n}, \\ &A^{bis}(m) = \max_{m'\in\mathcal{M}_n}\left\{\left\|\hat{h}_{m'}^{\hat{G}} - \hat{h}_{m\wedge m'}^{\hat{G}}\right\|^2 - V^{bis}(m')\right\}_+, \end{aligned} \tag{30}$$

avec c une constante numérique.
Notons que $\|\hat{h}_{m'}^{\hat{G}} - \hat{h}_{m\wedge m'}^{\hat{G}}\|^2 = \|\hat{f}_{m'}^{\hat{G},G} - \hat{f}_{m\wedge m'}^{\hat{G},G}\|_g^2$.
On sélectionne ensuite,

$$\hat{m}^{\hat{G},l} = \mathrm{argmin}_{m\in\mathcal{M}_n}\{A^{bis}(m) + 2V^{bis}(m)\}. \tag{31}$$

L'estimateur sélectionné ainsi est donc $\hat{f}_{\hat{m}^{\hat{G},l}}^{\hat{G},\hat{G},l}$.

9.2. **Construction de l'estimateur par pénalisation.**

On définit l'application pen$'$:

$$m \in \mathcal{M}_n \quad \mapsto \quad \mathrm{pen}'(m) := c\frac{D_m}{n}, \tag{32}$$

avec c une constante dépendant de $\mathbb{E}[Y_1^2]$ et de $\|h'\|$. Soit aussi $\hat{m}^{\hat{G},p} \in \mathrm{argmin}_{m\in\mathcal{M}_n}\gamma_n^{\hat{G}}(\hat{h}_m^{\hat{G}}) + \mathrm{pen}'(m)$.

L'estimateur sélectionné est $\hat{f}_{\hat{m}^{\hat{G},p}}^{\hat{G},\hat{G},p} = \hat{h}_{\hat{m}}^{\hat{G}} \circ \hat{G}_n$.

9.3. **Majoration du risque pour chacun des deux estimateurs.**

On va obtenir les deux théorèmes ci-dessous, sous les hypothèses suivantes :
- on se place sur les modèles trigonométriques,
- les dimensions des modèles vérifient $D_m \leq n^{1/3}/\ln(n)$ (quel que soit $m \in \mathcal{M}_n$),
- la fonction h est de classe $\mathcal{C}^1$ sur $[0;1]$ (donc bornée) et dans $W^{per}(1,L)$,
- le bruit ε_1 admet un moment d'ordre $2+p$, $p > 8/3$,
- $n \geq n_0 = \exp(\|h'\|^2)$.

Théorème 5. *L'estimateur sélectionné par méthode de Lepski vérifie l'inégalité oracle suivante :*

$$\begin{aligned} \mathbb{E}\left[\left\|\hat{f}_{\hat{m}^{\hat{G},l}}^{\hat{G},\hat{G},l} - f\right\|_g^2\right] \quad \leq \quad & \min_{m\in\mathcal{M}_n}\left(18V^{bis}(m) + 27\left(\bar{C} + \phi_0^2\mathbb{E}\left[Y_1^2\right]\right)\frac{D_m}{n}\right. \\ & \left.+217\left\|f - f_m^G\right\|_g^2\right) + \frac{C\ln(n)}{n}, \end{aligned}$$

où l'on rappelle que $f_m^G = h_m \circ G$, h_m *étant la projection orthogonale de* h *sur* S_m, *et avec* C *une constante dépendant de* L, ϕ_0^2, $\|h\|_\infty$, $\|h\|$, $\|h'\|$, $\|\varphi_2^{(r)}\|_\infty$ *pour* $r = 1, 2$, C_p *et* δ, *et avec*

$$\bar{C} = 6\left(C_2\|h'\|^2 + \phi_0^2\mathbb{E}[Y_1^2]\right) + 3C_4(\pi^4/4)\|h\|^2 + \frac{C_6}{2}\left\|\varphi_2^{(3)}\right\|_\infty^2 \|h\|^2.$$

Théorème 6. *L'estimateur sélectionné par pénalisation vérifie l'inégalité-oracle suivante :*

$$(33) \qquad \mathbb{E}\left[\left\|\hat{f}_{\hat{m}^{\hat{G},p}}^{\hat{G},\hat{G},p} - f\right\|_g^2\right] \leq \min_{m\in\mathcal{M}_n}\left\{3\frac{\theta+2}{\theta-2}\left\|f_m^G - f\right\|_g^2 + \frac{6\theta}{\theta-2}pen'(m)\right\} + \frac{C\ln(n)}{n},$$

où l'on rappelle que $f_m^G = h_m \circ G$, h_m *étant la projection orthogonale de* h *sur* S_m, *et où* C *est une constante ne dépendant que de* ϕ_0^2, δ, L, $\|f\|_\infty$, $\|h\|$, $\|h'\|$, $\|\varphi_2^{(r)}\|_{\infty,[0;1]}$, *pour* $r = 1, 2$ *des constantes* C_p, $p = 2, 3, 4, 6$ *de l'inégalité de Dvoretzky-Kiefer-Wolfowitz.*

Corollaire 25. *Avec les notations et hypothèses des théorèmes, en supposant* $h \in W^{per}(\alpha, L)$ *(à la place de* $W^{per}(1, L)$*), pour* $\alpha \geq 1$, *et pour* $\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}$ *l'un des deux estimateurs étudiés,*

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq Cn^{\frac{-2\alpha}{2\alpha+1}}.$$

9.4. Résultats préliminaires aux preuves des Théorèmes 6 et 5.

On commence par énoncer des lemmes utiles à la preuve. On utilise les notations de la section 8.3.1 et plus précisément de la preuve de la Proposition 17 : il s'agit de recentrer convenablement les termes T_i^m, $i = 1, \ldots, 4$, $m \in \mathcal{M}_n$ de manière à rendre négligeable le maximum en m de leur partie positive.

Lemme 26. *Sous les hypothèses du Théorème 5, il existe une constante numérique* C, *dépendant de* ϕ_0 *et de* $\|\varphi_2'\|_\infty$, *telle que*

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_2^{m'} - V_2(m')\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec pour $\kappa' = 7/3$ *et* $\kappa = 96\phi_0^2\mathbb{E}[Y_1^2]\|\varphi_2'\|_\infty^2$,

$$V_2(m') = \kappa\kappa'\frac{D_{m'}^4}{n^2}\ln^2(n).$$

Sous l'hypothèse $D_{m'} \leq (n/\ln(n)^2)^{1/3}$, on a

$$V_2(m') \leq \kappa\kappa'\frac{D_{m'}}{n} := V_2^{bis}(m').$$

L'inégalité du lemme est encore valable pour V_2 remplacé par V_2^{bis}.

Lemme 27. *Sous les hypothèses du Théorème 5, l'inégalité suivante est valide, pour* $p_{m'} = m'$ *ou* $m \wedge m'$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{p_{m'},b} - V_3(m')\right)_+\right] \leq \frac{C}{n}.$$

avec $V_3(m') = c_3\frac{D_{m'}}{n}$, *et* c_3 *une constante numérique, dépendant des quantités inconnues seulement au travers de* $\mathbb{E}[Y_1^2]$.

Lemme 28. *Sous les hypothèses du Théorème 5, l'inégalité suivante est valide, pour* $n \geq n_0$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_4^{m'} - V_4(m')\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec $V_4(m') = c_4\frac{D_{m'}}{n}$*,* c_4 *une constante numérique, dépendant des quantités inconnues seulement au travers de* $\mathbb{E}[Y_1^2]$ *et*

$$n_0 = \exp\left(\|h'\|^2\right).$$

On peut de plus remplacer V_4 *par une quantité* $\tilde{V}_4$ *dépendant de* $\|h'\|^2$*. Le résultat est alors valable quel que soit l'entier* n*.*

9.5. **Preuve du Théorème 5.**

Dans toute la preuve C désignera une constante dépendant uniquement des quantités indiquées dans l'énoncé du théorème et pouvant varier d'une ligne à l'autre. Pour simplifier, on abrège A^{bis} en A, $\hat{m}^{\hat{G},l}$ en $\hat{m}$ et $\hat{f}_{\hat{m}}^{\hat{G},\hat{G},l}$ en $\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}$.
On fixe $m \in \mathcal{M}_n$. On commence par majorer

$$\begin{aligned}
\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2 &\leq 3\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - \hat{f}_{\hat{m}}^{\hat{G},G} - \mathbb{E}\left[\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - \hat{f}_{\hat{m}}^{\hat{G},G}|(X_{-l})_l\right]\right\|_g^2 \\
&\quad +3\left\|\mathbb{E}\left[\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - \hat{f}_{\hat{m}}^{\hat{G},G}|(X_{-l})_l\right]\right\|_g^2 + 3\left\|\hat{f}_{\hat{m}}^{\hat{G},G} - f\right\|_g^2, \\
&= 3T_2^{\hat{m}} + 3T_4^{\hat{m}} + 3\left\|\hat{h}_{\hat{m}}^{\hat{G}} - h\right\|^2,
\end{aligned}$$

en utilisant les notations de la preuve de la Proposition 17. On poursuit ensuite en faisant apparaître A et V^{bis}, et en se débarrassant de l'indice aléatoire $\hat{m}$ pour le dernier terme de la majoration ci-dessus :

$$\begin{aligned}
\left\|\hat{h}_{\hat{m}}^{\hat{G}} - h\right\|^2 &\leq 3\left\|\hat{h}_{\hat{m}}^{\hat{G}} - \hat{h}_{m\wedge\hat{m}}^{\hat{G}}\right\|^2 + 3\left\|\hat{h}_{m\wedge\hat{m}}^{\hat{G}} - \hat{h}_{m}^{\hat{G}}\right\|^2 + 3\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2, \\
&\leq 3\left(A(m) + V^{bis}(\hat{m})\right) + 3\left(A(\hat{m}) + V^{bis}(\hat{m})\right) + 3\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2, \\
&= 3\left(A(m) + 2V^{bis}(m)\right) + 3\left(A(\hat{m}) + 2V^{bis}(\hat{m})\right) + 3\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2 - 3V^{bis}(\hat{m}) - 3V^{bis}(m), \\
&\leq 3\left(A(m) + 2V^{bis}(m)\right) + 3\left(A(\hat{m}) + 2V^{bis}(\hat{m})\right) + 3\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2 - 2V^{bis}(\hat{m}), \\
&\leq 6\left(A(m) + 2V^{bis}(m)\right) - 2V^{bis}(\hat{m}) + 3\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2,
\end{aligned}$$

par définition de $\hat{m}$. Il faut donc maintenant majorer $\|\hat{h}_m^{\hat{G}} - h\|^2$, qui n'est pas, contrairement à ce qui se passait dans le cas "G connue", la perte de l'estimateur étudié en non-adaptatif. On se ramène cependant à des quantités déjà étudiées :

$$\begin{aligned}
\left\|\hat{h}_{m}^{\hat{G}} - h\right\|^2 &= \left\|\hat{f}_{m}^{\hat{G},G} - f\right\|_g^2, \\
&\leq 3\left\|\hat{f}_{m}^{\hat{G},G} - \hat{f}_{m}^{G,G} - \mathbb{E}\left[\hat{f}_{m}^{\hat{G},G} - \hat{f}_{m}^{G,G}|(X_{-l})_l\right]\right\|_g^2 \\
&\quad +3\left\|\mathbb{E}\left[\hat{f}_{m}^{\hat{G},G} - \hat{f}_{m}^{G,G}|(X_{-l})_l\right]\right\|_g^2 + 3\left\|\hat{f}_{m}^{G,G} - f\right\|_g^2, \\
&= 3T_1^m + 3T_3^m + 3T_0^m,
\end{aligned}$$

en reprenant toujours les notations introduites dans la preuve de la Proposition 17. Par suite, on obtient

$$\begin{aligned}\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}-f\right\|_g^2 &\leq 3T_2^{\hat{m}}+3T_4^{\hat{m}}-3\times 2V^{bis}(\hat{m})+3\times 6\Big(A(m)+V^{bis}(m)\Big)\\ &\quad +3\times 3\times\left(3T_1^m+3T_3^m+3T_0^m\right).\end{aligned}$$

Or, on a démontré en non-adaptatif les inégalités suivantes :

$$\begin{aligned}\mathbb{E}\left[T_0^m\right] &\leq \phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m}{n}+\left\|f-f_m^G\right\|_g^2,\\ \mathbb{E}\left[T_1^m\right] &\leq \frac{C}{n},\ \text{pour } D_m\leq n^{1/3},\\ \mathbb{E}\left[T_3^m\right] &\leq \bar{C}\frac{D_m}{n}+\frac{C}{n},\ \text{pour } D_m\leq n^{1/3},\end{aligned}$$

avec

$$\bar{C}=6\left(C_2\|h'\|^2+\phi_0^2\mathbb{E}[Y_1^2]\right)+3C_4(\pi^4/4)\|h\|^2+\frac{C_6}{2}\left\|\varphi_2^{(3)}\right\|_\infty^2\|h\|^2.$$

On obtient ainsi provisoirement :

$$\begin{aligned}\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}-f\right\|_g^2\right] &\leq 18\Big(\mathbb{E}\left[A(m)\right]+V^{bis}(m)\Big)+3\mathbb{E}\left[\Big(T_2^{\hat{m}}-V^{bis}(\hat{m})\Big)_+\right]+3\mathbb{E}\left[\Big(T_4^{\hat{m}}-V^{bis}(\hat{m})\Big)_+\right]\\ &\quad +27\left(\bar{C}+\phi_0^2\mathbb{E}\left[Y_1^2\right]\right)\frac{D_m}{n}+\left\|f-f_m^G\right\|_g^2+\frac{C}{n}.\end{aligned}$$

Il reste à majorer $A(m)$, $T_2^{\hat{m}}$ et $T_4^{\hat{m}}$. Concernant ces deux derniers termes, on commence par les majorer de la façon suivante : pour $k=2,4$,

$$\mathbb{E}\left[\Big(T_k^{\hat{m}}-V^{bis}(\hat{m})\Big)_+\right]\leq\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\Big(T_k^{m'}-V^{bis}(m')\Big)_+\right],$$

puis on utilise les Lemmes 26 et 28, en notant que par définition, $V^{bis}\geq V_2^{bis}$ et $V^{bis}\geq V_4$. On obtient donc

$$\begin{aligned}\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}-f\right\|_g^2\right] &\leq 18\Big(\mathbb{E}\left[A(m)\right]+V^{bis}(m)\Big)\\ &\quad +27\left(\bar{C}+\phi_0^2\mathbb{E}\left[Y_1^2\right]\right)\frac{D_m}{n}\\ &\quad +\left\|f-f_m^G\right\|_g^2+\frac{C\ln(n)}{n}.\end{aligned}$$

La conclusion du théorème provient donc du lemme suivant :

Lemme 29. *Sous les hypothèses du théorème, il existe une constante $C\geq 0$, telle que, pour tout $m\in\mathcal{M}_n$,*

$$\mathbb{E}\left[A(m)\right]\leq C\frac{1}{n}+12\left\|f_m^G-f\right\|_g^2.$$

□

9.6. **Preuve du Théorème 6.**

On notera dans la suite C une constante pouvant varier d'une ligne à l'autre. Pour simplifier, on abrège $\hat{m}^{\hat{G},p}$ en $\hat{m}$ et $\hat{f}_{\hat{m}}^{\hat{G},\hat{G},p}$ en $\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}$.
Soit $m\in\mathcal{M}_n$ fixé. Par définition de $\hat{h}_{\hat{m}}$, on a

$$\gamma_n^{\hat{G}}(\hat{h}_{\hat{m}})-\gamma_n^{\hat{G}}(h_m)\leq \mathrm{pen}'(m)-\mathrm{pen}'(\hat{m}).$$

On utilise ensuite la définition du contraste, ainsi que la décomposition, pour $t \in L^2([0;1])$,

$$\|t\|^2 = \|t \circ G - f\|_g^2 - \|f\|_g^2 + 2\langle t \circ G, f\rangle_g,$$

pour en déduire

$$\left\| \hat{f}_{\hat{m}}^{\hat{G},G} - f \right\|_g^2 \leq \left\| f_m^G - f \right\|_g^2 + \mathrm{pen}'(m) - \mathrm{pen}'(\hat{m}) + 2\tilde{\nu}_n(\hat{h}_{\hat{m}}^{\hat{G}} - h_m),$$

où l'on a défini, pour $t \in L^2([0;1]$,

$$\tilde{\nu}_n(t) = \frac{1}{n}\sum_{i=1}^n Y_i t\left(\hat{G}_n(X_i)\right) - \langle t \circ G, f\rangle_g,$$

processus empirique non centré. On le majore de la façon suivante :

$$\begin{aligned} 2\tilde{\nu}_n\left(\hat{h}_{\hat{m}}^{\hat{G}} - h_m\right) &= 2\left\|\hat{h}_{\hat{m}}^{\hat{G}} - h_m\right\| \tilde{\nu}_n\left(\frac{\hat{h}_{\hat{m}}^{\hat{G}} - h_m}{\left\|\hat{h}_{\hat{m}}^{\hat{G}} - h_m\right\|}\right) \\ &\leq 2\left\|\hat{h}_{\hat{m}}^{\hat{G}} - h_m\right\| \sup_{\substack{t \in S_m + S_{\hat{m}} \\ \|t\|=1}} |\tilde{\nu}_n(t)|\,. \end{aligned}$$

On introduit ensuite le paramètre $\theta > 0$, et on a,

$$2\tilde{\nu}_n(\hat{h}_{\hat{m}}^{\hat{G}} - h_m) \leq \frac{2}{\theta}\left(\left\|\hat{f}_{\hat{m}}^{\hat{G},G} - f\right\|_g^2 + \left\|f_m^G - f\right\|_g^2\right) + \theta \sup_{t \in \mathcal{S}(m \vee \hat{m})} (\tilde{\nu}_n(t))^2,$$

avec $\mathcal{S}(m \vee \hat{m}) = \{t \in S_m + S_{\hat{m}},\ \|t\| = 1\}$. Ceci entraîne,

$$\frac{\theta-2}{\theta}\left\|\hat{f}_{\hat{m}}^{\hat{G},G} - f\right\|_g^2 \leq \frac{\theta+2}{\theta}\left\|f_m^G - f\right\|_g^2 + \theta \sup_{t \in \mathcal{S}(m \vee \hat{m})} (\tilde{\nu}_n(t))^2 + \mathrm{pen}'(m) - \mathrm{pen}'(\hat{m}). \tag{34}$$

Or, on souhaite contrôler le risque de l'estimateur $\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}$ et pas celui de $\hat{f}_{\hat{m}}^{\hat{G},G}$. On utilise donc,

$$\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2 \leq 3\left\|\hat{f}_{\hat{m}}^{\hat{G},G} - f\right\|_g^2 + 3T_2^{\hat{m}} + 3T_4^{\hat{m}},$$

où l'on a noté

$$\begin{aligned} T_2^{\hat{m}} &= \left\|\hat{f}_{\hat{m}}^{\hat{G},G} - \hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - \mathbb{E}\left[\hat{f}_{\hat{m}}^{\hat{G},G} - \hat{f}_{\hat{m}}^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g^2, \\ T_4^{\hat{m}} &= \left\|\mathbb{E}\left[\hat{f}_{\hat{m}}^{\hat{G},G} - \hat{f}_{\hat{m}}^{\hat{G},\hat{G}} |(X_{-l})_l\right]\right\|_g^2, \end{aligned}$$

pour reprendre des notations utilisées dans l'étude non adaptative. Introduisant ceci dans (34), on obtient :

$$\frac{\theta-2}{\theta}\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2 \leq 3\frac{\theta+2}{\theta}\left\|f_m^G - f\right\|_g^2 + 3\theta \sup_{t \in \mathcal{S}(m \vee \hat{m})} (\tilde{\nu}_n(t))^2 + 3(\mathrm{pen}'(m) - \mathrm{pen}'(\hat{m})) + 3T_2^{\hat{m}} + 3T_4^{\hat{m}}.$$

Le nouveau processus $\tilde{\nu}_n$ n'étant pas centré, on se ramène au précédent ν_n comme dans la partie précédente :

$$\sup_{t \in \mathcal{S}(m \vee \hat{m})} (\tilde{\nu}_n(t))^2 \leq 2 \sup_{t \in \mathcal{S}(m \vee \hat{m})} (\nu_n(t))^2 + 2T_{sup}^{\hat{m} \vee m},$$

avec

$$T_{sup}^{\hat{m} \vee m} = \sup_{t \in \mathcal{S}(m \vee \hat{m})} \left(\frac{1}{n}\sum_{i=1}^n t\left(\hat{G}_n(X_i)\right) Y_i - t\left(G(X_i)\right) Y_i\right)^2.$$

On majore de même ce terme :

$$T_{sup}^{\hat{m}\vee m} \leq 2T_1^{\hat{m}\vee m,b} + 2T_3^{\hat{m}\vee m,b}$$

avec les notations de la partie non-adaptative :

$$\begin{aligned}
T_1^{m\vee\hat{m},b} &= \sum_{j=1}^{D_{m\vee\hat{m}}} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l\right]\right)^2, \\
T_3^{m\vee\hat{m},b} &= \sum_{j=1}^{D_{m\vee\hat{m}}} \left(\mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l\right]\right)^2.
\end{aligned}$$

Par suite, pour $\theta > 2$,

$$\begin{aligned}
\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2\right] \leq\; & 3\frac{\theta+2}{\theta-2}\left\|f_m^G - f\right\|_g^2 + \frac{3\theta}{\theta-2}\left(\mathrm{pen}'(m) - \mathbb{E}\left[\mathrm{pen}'(\hat{m})\right]\right) \\
& + \frac{6\theta^2}{\theta-2}\mathbb{E}\left[\sup_{t\in\mathcal{S}(m\vee\hat{m})} (\nu_n(t))^2\right] \\
& + \frac{3\theta}{\theta-2}\left(\mathbb{E}\left[T_2^{\hat{m}}\right] + \mathbb{E}[T_4^{\hat{m}}] + 4\left(\mathbb{E}\left[T_1^{\hat{m}\vee m,b}\right] + \mathbb{E}\left[T_3^{\hat{m}\vee m,b}\right]\right)\right).
\end{aligned}$$

On retranche maintenant aux quantités de cette majorations, les termes $p(m,\hat{m})$ de la preuve du Théorème 3, V_i (ou $\tilde{V}_i$, ou V_i^{bis}) $i = 2,3,4$ intervenant dans la preuve du Théorème 5 (le terme non défini, $\tilde{V}_4$, étant le terme V_4 "non asymptotique", c'est-à-dire obtenu avant de s'être affranchi de la dépendance en $\|h'\|$, cf. la preuve du Lemme 28) :

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2\right]$$

$$\begin{aligned}
\text{(35)} \qquad \leq\; & 3\frac{\theta+2}{\theta-2}\left\|f_m^G - f\right\|_g^2 + \frac{3\theta}{\theta-2}\left(\mathrm{pen}'(m) - \mathbb{E}\left[\mathrm{pen}'(\hat{m})\right]\right) \\
& + \frac{6\theta^2}{\theta-2}\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m\vee\hat{m})} (\nu_n(t))^2 - p(m,\hat{m})\right)_+\right] \\
& + \frac{3\theta}{\theta-2}\left(\mathbb{E}\left[\left(T_2^{\hat{m}} - V_2(\hat{m})\right)_+\right] + \mathbb{E}\left[\left(T_4^{\hat{m}} - \tilde{V}_4(\hat{m})\right)_+\right],\right. \\
& \left. + 4\left\{\mathbb{E}\left[T_1^{\hat{m}\vee m,b}\right] + \mathbb{E}\left[\left(T_3^{\hat{m}\vee m,b} - V_3(\hat{m}\vee m)\right)_+\right]\right\}\right) \\
& + \frac{6\theta^2}{\theta-2}\mathbb{E}\left[p(m,\hat{m})\right] + \frac{3\theta}{\theta-2}\mathbb{E}\left[V_2(\hat{m}) + \tilde{V}_4(\hat{m}) + 4V_3(\hat{m}\vee m)\right].
\end{aligned}$$

On obtient donc encore la majoration :

$$\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}} - f\right\|_g^2\right]$$

$$\begin{aligned}
\leq\ & 3\frac{\theta+2}{\theta-2}\left\|f_m^G-f\right\|_g^2+\frac{3\theta}{\theta-2}\left(\mathrm{pen}'(m)-\mathbb{E}\left[\mathrm{pen}'(\hat{m})\right]\right)\\
&+\frac{6\theta^2}{\theta-2}\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m\vee\hat{m})}(\nu_n(t))^2-p(m,\hat{m})\right)_+\right]\\
&+\frac{3\theta}{\theta-2}\left(\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_2^{m'}-V_2^{bis}(m')\right)_+\right]+\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_4^{m'}-\tilde{V}_4(m')\right)_+\right],\right.\\
&\left.+4\left\{\mathbb{E}\left[T_1^{\hat{m}\vee m,b}\right]+\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{m',b}-V_3(m')\right)_+\right]\right\}\right)\\
&+\frac{6\theta^2}{\theta-2}\mathbb{E}\left[p(m,\hat{m})\right]+\frac{3\theta}{\theta-2}\mathbb{E}\left[V_2^{bis}(\hat{m})+\tilde{V}_4(\hat{m})+4V_3(\hat{m}\vee m)\right],
\end{aligned}$$

à laquelle on peut appliquer les Lemmes 7, 26, 28 et 27 pour obtenir

$$\begin{aligned}
\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}-f\right\|_g^2\right] \leq\ & 3\frac{\theta+2}{\theta-2}\left\|f_m^G-f\right\|_g^2\\
&+\frac{3\theta}{\theta-2}\left(\mathrm{pen}'(m)-\mathbb{E}\left[\mathrm{pen}'(\hat{m})\right]\right)+4\mathbb{E}\left[T_1^{\hat{m}\vee m,b}\right]+\frac{C\ln(n)}{n}\\
&+\frac{6\theta^2}{\theta-2}\mathbb{E}\left[p(m,\hat{m})\right]+\frac{3\theta}{\theta-2}\mathbb{E}\left[V_2^{bis}(\hat{m})+\tilde{V}_4(\hat{m})+4V_3(\hat{m}\vee m)\right].
\end{aligned}$$

On majore $T_1^{\hat{m}\vee m,b}$ de la façon suivante

$$\begin{aligned}
\mathbb{E}\left[T_1^{m\vee\hat{m},b}\right] &= \sum_{j=1}^{D_{m\vee\hat{m}}}\left(\hat{a}_j^{\hat{G}}-\hat{a}_j^{G}-\mathbb{E}\left[\hat{a}_j^{\hat{G}}-\hat{a}_j^{G}\,|(X_{-l})_l\right]\right)^2,\\
&\leq \sum_{j=1}^{D_{m_{\max}}}\left(\hat{a}_j^{\hat{G}}-\hat{a}_j^{G}-\mathbb{E}\left[\hat{a}_j^{\hat{G}}-\hat{a}_j^{G}\,|(X_{-l})_l\right]\right)^2,\\
&\leq C\frac{D_{m_{\max}}^3}{n^2}\leq C\frac{1}{n},
\end{aligned}$$

sous la condition $D_{m_{\max}}\leq n^{1/3}$. Par suite,

$$\begin{aligned}
\mathbb{E}\left[\left\|\hat{f}_{\hat{m}}^{\hat{G},\hat{G}}-f\right\|_g^2\right] \leq\ & 3\frac{\theta+2}{\theta-2}\left\|f_m^G-f\right\|_g^2+\frac{3\theta}{\theta-2}\mathrm{pen}'(m)+\frac{C\ln(n)}{n},\\
&+\frac{3\theta}{\theta-2}\mathbb{E}\left[2\theta p(m,\hat{m})+V_2^{bis}(\hat{m})+\tilde{V}_4(\hat{m})+4V_3(\hat{m}\vee m)-\mathrm{pen}'(\hat{m})\right].
\end{aligned}$$

On rappelle enfin que

$$\begin{aligned}
&p(m,\hat{m})=6(1+2\delta)\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_m+D_{\hat{m}}}{n},\\
&V_2^{bis}(\hat{m})=c_2\frac{D_{\hat{m}}}{n},V_2^{bis}(\hat{m})=c_2\frac{D_{\hat{m}}}{n},\\
&V_3(m\vee\hat{m})=c_3\frac{D_{m\vee\hat{m}}}{n},\\
&\tilde{V}_4(\hat{m})=c_4\frac{D_{\hat{m}}}{n},
\end{aligned}$$

avec c_i des constantes numériques dépendant (linéairement) de $\mathbb{E}[Y_1^2]$, et c_4 dépendant (linéairement) de $\|h'\|$. On peut donc majorer, pour $i = 2, 4, 3$, les quantités A_i, avec $A_2 = V_2^{bis}(\hat{m})$, $A_3 = V_3(m \vee \hat{m})$, et $A_4 = \tilde{V}_4(\hat{m})$.

$$A_i \leq c\frac{D_m + D_{\hat{m}}}{n},$$

le même majorant étant valable pour $p(m, \hat{m})$. En choisissant

$$\text{pen}'(m) = c\frac{D_m}{n},$$

on a donc

$$\mathbb{E}\left[2\theta p(m, \hat{m}) + V_2^{bis}(\hat{m}) + \tilde{V}_4(\hat{m}) + 4V_3(\hat{m} \vee m) - \text{pen}'(\hat{m})\right] \leq \text{pen}'(m),$$

et l'inégalité oracle en découle.

□

9.7. **Preuve du Lemme 29.**

Soit $m' \in \mathcal{M}_n$. On décompose,

$$\left\|\hat{h}_{m'}^{\hat{G}} - \hat{h}_{m\wedge m'}^{\hat{G}}\right\|^2 \leq 3\left\|\hat{h}_{m'}^{\hat{G}} - h_{m'}\right\|^2 + 3\left\|h_{m'} - h_{m\wedge m'}\right\|^2 + 3\left\|h_{m\wedge m'} - \hat{h}_{m\wedge m'}^{\hat{G}}\right\|^2.$$

On note ensuite que, pour $p = m'$ ou $p = m \wedge m'$ $\|h_p - \hat{h}_p^{\hat{G}}\|^2 = \sup_{t\in\mathcal{S}(p)} \tilde{\nu}_n(t)^2$, où $\mathcal{S}(p)$ désigne la sphère unité de S_p, et où l'on a défini

$$\tilde{\nu}_n(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ \hat{G}_n(X_i) - \mathbb{E}\left[Y_i t \circ G(X_i)\right].$$

En effet, si $t \in \mathcal{S}(p)$, $t = \sum_{j=1}^{D_p} b_j\varphi_j$, avec $\sum_j b_j^2 = 1$. On peut donc écrire :

$$\tilde{\nu}_n(t)^2 \quad \leq \quad \left(\sum_{j=1}^{D_p} b_j\tilde{\nu}_n(\varphi_j)\right)^2 = \sum_{j=1}^{D_p} \tilde{\nu}_n^2(\varphi_j)$$

par l'inégalité de Cauchy-Schwarz, avec égalité pour le choix $b_j = \tilde{\nu}_n(\varphi_j)/\sqrt{\sum_j \tilde{\nu}_n^2(\varphi_j)}$. Il reste ensuite à remarquer que $\tilde{\nu}_n(\varphi_j) = \hat{a}_j^{\hat{G}} - a_j$, ce qui donne, en élevant au carré et en sommant la norme $\|\hat{h}_p^{\hat{G}} - h_p\|^2$. On se ramène ensuite au processus ν_n centré de la façon suivante : quel que soit $t \in L^2([0; 1])$,

$$\begin{aligned}
\tilde{\nu}_n(t) &= \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ \hat{G}_n(X_i) - \langle t \circ G, f\rangle_g, \\
&= \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ G(X_i) - \langle t \circ G, f\rangle_g + \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ (\hat{G}_n(X_i) - G(X_i)), \\
&= \nu_n(t) + \frac{1}{n}\sum_{i=1}^{n} Y_i t \circ (\hat{G}_n(X_i) - G(X_i)),
\end{aligned}$$

et ainsi,

$$(\tilde{\nu}_n(t))^2 \leq 2\left(\nu_n(t)\right)^2 + 2\left(\frac{1}{n}\sum_{i=1}^{n} Y_i t \circ (\hat{G}_n(X_i) - G(X_i))\right)^2.$$

On utilise maintenant la remarque qui suit : $t \in S_p$ avec $\|t\| = 1$, si et seulement si il existe une famille de réels $\theta = (\theta_j)_{j=1,\ldots,D_p}$ telle que

$$t = \sum_{j=1}^{p} \theta_j \varphi_j, \quad \text{avec } \sum_{j=1}^{D_p} \theta_j^2 = 1.$$

Ceci permet d'écrire,

$$\begin{aligned}
&\sup_{t \in \mathcal{S}(p)} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i t \circ (\hat{G}_n(X_i) - G(X_i)) \right)^2 \\
&\quad = \sup_{\substack{\theta \in \mathbb{R}^p \\ \sum_j \theta_j^2 = 1}} \left(\sum_{j=1}^{D_p} \theta_j \frac{1}{n} \sum_{i=1}^{n} Y_i \varphi_j \circ (\hat{G}_n(X_i) - G(X_i)) \right)^2, \\
&\quad = \sup_{\substack{\theta \in \mathbb{R}^p \\ \sum_j \theta_j^2 = 1}} \left(\sum_{j=1}^{D_p} \theta_j \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \right) \right)^2, \\
&\quad = \sup_{\substack{\theta \in \mathbb{R}^p \\ \sum_j \theta_j^2 = 1}} \left(\sum_{j=1}^{D_p} \theta_j^2 \right) \left(\sum_{j=1}^{\hat{D}_p} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \right)^2 \right),
\end{aligned}$$

en effet, on a une majoration par l'inégalité de Cauchy-Schwarz, et l'égalité provient du choix

$$\forall j = 1, \ldots, D_p, \quad \theta_j = \frac{\hat{a}_j^{\hat{G}} - \hat{a}_j^{G}}{\sqrt{\sum_{j=1}^{D_p} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \right)^2}}.$$

On en déduit donc,

$$\begin{aligned}
&\sup_{t \in \mathcal{S}(p)} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i t \circ (\hat{G}_n(X_i) - G(X_i)) \right)^2 \\
&\quad = \sum_{j=1}^{D_p} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \right)^2, \\
&\quad \leq 2 \sum_{j=1}^{D_p} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l \right] \right)^2 + 2 \sum_{j=1}^{D_p} \left(\mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l \right] \right)^2, \\
&\quad := 2T_1^{p,b} + 2T_3^{p,b},
\end{aligned}$$

avec

$$\begin{aligned}
T_1^{p,b} &= \sum_{j=1}^{D_p} \left(\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} - \mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l \right] \right)^2, \\
T_3^{p,b} &= \sum_{j=1}^{D_p} \left(\mathbb{E}\left[\hat{a}_j^{\hat{G}} - \hat{a}_j^{G} \,|(X_{-l})_l \right] \right)^2.
\end{aligned}$$

Par conséquent,

$$\left\|h_p - \hat{h}_p^{\hat{G}}\right\|^2 \leq 2 \sup_{t\in\mathcal{S}(p)} (\nu_n(t))^2 + 4T_1^{p,b} + 4T_3^{p,b}.$$

On a donc

$$\begin{aligned}\left\|\hat{h}_{m'}^{\hat{G}} - \hat{h}_{m\wedge m'}^{\hat{G}}\right\|^2 &\leq 6 \sup_{t\in\mathcal{S}(m')} (\nu_n(t))^2 + 6 \sup_{t\in\mathcal{S}(m\wedge m')} (\nu_n(t))^2 + 12T_3^{m',b} + 12T_3^{m\wedge m',b} \\ &\quad +12T_1^{m',b} + 12T_1^{m\wedge m',b} + 3\left\|h_{m'} - h_{m\wedge m'}\right\|^2.\end{aligned}$$

Ce que l'on réécrit en retranchant $V(m')$ et en intégrant

$$\begin{aligned}\mathbb{E}\left[A(m)\right] &= \mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(\left\|\hat{h}_{m'}^{\hat{G}} - \hat{h}_{m\wedge m'}^{\hat{G}}\right\|^2 - V^{bis}(m')\right)_+\right] \\ &\leq 6\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m')}(\nu_n(t))^2 - \frac{V^{bis}(m')}{24}\right)_+\right] + 6\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m\wedge m')}(\nu_n(t))^2 - \frac{V^{bis}(m')}{24}\right)_+\right] \\ &\quad +12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{m',b} - \frac{V^{bis}(m')}{48}\right)_+\right] + 12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{m\wedge m',b} - \frac{V^{bis}(m')}{48}\right)_+\right] \\ &\quad +12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} T_1^{m',b}\right] + 12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} T_1^{m\wedge m',b}\right] + 3\max_{m'\in\mathcal{M}_n}\left\|h_{m'} - h_{m\wedge m'}\right\|^2.\end{aligned}$$

Le terme non aléatoire est T_c^m et a déjà été majoré (partie adaptation par méthode de Lepski dans le cas où la répartition G était supposée connue) :

$$\max_{m'\in\mathcal{M}_n}\left\|h_{m'} - h_{m\wedge m'}\right\|^2 \leq 4\left\|h_m - h\right\|^2.$$

De plus, on a montré les deux inégalités suivantes, quand on a traité l'adaptation par méthode de Lepski dans le cas où la loi de G était supposée connue (voir les termes T_a et T_b^m ci-dessus) :

$$\begin{aligned}\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m')}(\nu_n(t))^2 - V_0(m')\right)_+\right] &\leq \frac{C}{n}, \\ \mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(\sup_{t\in\mathcal{S}(m\wedge m')}(\nu_n(t))^2 - V_0(m')\right)_+\right] &\leq \frac{C}{n},\end{aligned}$$

avec la notation

$$V_0(m') = 6(1+2\delta)\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D_{m'}}{n}.$$

On obtient aussi pour $p = m'$ ou $p = m\wedge m'$ que

$$\begin{aligned}T_1^{p,b} &= \sum_{j=1}^{D_p}\left\{\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}}\right) - \mathbb{E}\left[\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right\}^2, \\ &\leq \sum_{j=1}^{D_{m_{\max}}}\left\{\left(\hat{a}_j^G - \hat{a}_j^{\hat{G}}\right) - \mathbb{E}\left[\hat{a}_j^G - \hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right\}^2, \\ &= T_1^{m_{\max}}.\end{aligned}$$

Par conséquent,

$$12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} T_1^{m',b}\right] + 12\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} T_1^{m\wedge m',b}\right] \leq 24\mathbb{E}\left[T_1^{m_{\max}}\right] \leq \frac{C}{n},$$

pour $D_{m_{\max}} \leq n^{1/3}$. On utilise enfin le Lemme 27. Pour conclure, on effectue le choix de V^{bis} (précisément le choix de la constante c intervenant dans sa définition) de telle sorte que

$$V^{bis} \geq 24V_0, \ V^{bis} \geq 48V_3.$$

On conclut donc

$$\mathbb{E}\left[A(m)\right] \leq C\frac{\ln(n)}{n} + 12\left\|h_m - h\right\|^2.$$

□

9.7.1. *Preuve du Lemme 26.*
On commence par

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_2^{m'} - V_2(m')\right)_+\right] \leq \sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(T_2^{m'} - V_2(m')\right)_+\right],$$

et on raisonne pour chaque $m' \in \mathcal{M}_n$: en utilisant la majoration (19) de la preuve du Lemme 22, on a d'abord,

$$\begin{aligned}\mathbb{E}\left[\left(T_2^{m'} - V_2(m')\right)_+\right] &\leq D_{m'}^3\left\|\varphi_2'\right\|_\infty^2\mathbb{E}\left[\left(\sum_{j=1}^{D_{m'}}\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2\left\|\hat{U}_n - id\right\|_\infty^2\right.\right.\\ &\quad\left.\left.-\frac{\kappa\kappa'}{\left\|\varphi_2'\right\|_\infty^2}\frac{D_{m'}}{n^2}\ln^2(n)\right)_+\right],\\ &\leq T_{2,a}^{m'} + T_{2,b}^{m'},\end{aligned}$$

où l'on note

$$\begin{aligned}T_{2,a}^{m'} &= D_{m'}^3\left\|\varphi_2'\right\|_\infty^2\mathbb{E}\left[\sum_{j=1}^{D_{m'}}\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa'\frac{\ln(n)}{n}\right)_+\right],\\ T_{2,b}^{m'} &= D_{m'}^3\left\|\varphi_2'\right\|_\infty^2\kappa'\frac{\ln(n)}{n}\mathbb{E}\left[\left(\sum_{j=1}^{D_{m'}}\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2 - \frac{\kappa}{\left\|\varphi_2'\right\|_\infty^2}\frac{D_{m'}}{n}\ln(n)\right)_+\right].\end{aligned}$$

Pour majorer $T_{2,a}^{m'}$, on commence par utiliser l'inégalité de Cauchy-Schwarz,

$$T_{2,a}^{m'} \leq D_{m'}^3\left\|\varphi_2'\right\|_\infty^2\sum_{j=1}^{D_{m'}}\mathbb{E}\left[\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^4\right]^{1/2}\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa'\frac{\ln(n)}{n}\right)^2\right]^{1/2}.$$

On majore ensuite chacun des deux termes principaux du produit :

$$\begin{aligned}
\mathbb{E}\left[\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^4\right] &\leq 2^{4-1}\mathbb{E}\left[\left(\hat{a}_j^{\hat{G}}\right)^4 + \left(\mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^4\right], \\
&\leq 8\mathbb{E}\left[\left(\hat{a}_j^{\hat{G}}\right)^4 + \mathbb{E}\left[\left(\hat{a}_j^{\hat{G}}\right)^4 |(X_{-l})_l\right]\right], \\
&= 16\mathbb{E}\left[\left(\hat{a}_j^{\hat{G}}\right)^4\right], \\
&= 16\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n Y_i\varphi_j \circ \hat{G}_n(X_i)\right)^4\right], \\
&\leq 16\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\left(Y_i\varphi_j \circ \hat{G}_n(X_i)\right)^4\right], \\
&\leq 16\mathbb{E}\left[Y_1^4\right]\|\varphi_j\|_\infty^4 \leq 16\phi_0^4\mathbb{E}\left[Y_1^4\right].
\end{aligned}$$

Ceci entraîne :

$$\sum_{j=1}^{D_{m'}} \mathbb{E}\left[\left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^4\right]^{1/2} \leq 4\phi_0^2\mathbb{E}\left[Y_1^4\right]^{1/2} D_{m'}.$$

D'autre part, la seconde inégalité du Corollaire 15 implique,

$$\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^2 - \kappa'\frac{\ln(n)}{n}\right)^2\right] \leq Cn^{-2-2\kappa'},$$

et donc en regroupant les deux dernières majorations,

$$\sum_{m'\in\mathcal{M}_n} T_{2,a}^{m'} \quad\leq\quad C\sum_{m'\in\mathcal{M}_n} D_{m'}^4 n^{-1-\kappa'} \leq Cn^{4/3-\kappa'} \leq Cn^{-1}$$

dès que $D_{m'} \leq n^{1/3}$ et pour la valeur $\kappa' = 7/3$.

Pour le terme $T_{2,b}^{m'}$, on commence par noter que

$$\sum_{j=1}^{D_{m'}} \left(\hat{a}_j^{\hat{G}} - \mathbb{E}\left[\hat{a}_j^{\hat{G}}\,|(X_{-l})_l\right]\right)^2 = \sum_{j=1}^{D_{m'}} \bar{\nu}_n^2(\varphi_j),$$

où l'on définit, pour $t \in L^2([0;1])$,

$$\bar{\nu}_n(t) = \frac{1}{n}\sum_{i=1}^n Y_i t\left(\hat{G}_n(X_i)\right) - \mathbb{E}\left[Y_i t\left(\hat{G}_n(X_i)\right)|(X_{-l})_l\right],$$

processus empirique centré conditionnellement à l'échantillon $(X_{-l})_l$. Un raisonnement déjà effectué donne ensuite

$$\sum_{j=1}^{D_{m'}} \bar{\nu}_n^2(\varphi_j) = \sup_{t\in\mathcal{S}(m')} \bar{\nu}_n^2(t).$$

Ce processus n'est cependant pas borné, on ne peut donc pas directement utiliser une inégalié de concentration. On effectue la décomposition suivante :

$$\bar{\nu}_n(t) = \bar{\nu}_n^{(1)}(t) + \bar{\nu}_n^{(2,1)}(t) + \bar{\nu}_n^{(2,2)}(t),$$

avec

$$\bar{\nu}_n^{(1)}(t) = \frac{1}{n}\sum_{i=1}^{n} f(X_i)t \circ \hat{G}_n(X_i) - \mathbb{E}\left[f(X_i)t\left(\hat{G}_n(X_i)\right)|(X_{-l})_l\right],$$

$$\bar{\nu}_n^{(2,1)}(t) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{1}_{|\varepsilon_i|\leq\kappa_n} t \circ \hat{G}_n - \mathbb{E}\left[\varepsilon_i \mathbf{1}_{|\varepsilon_i|\leq\kappa_n} t \circ \hat{G}_n \,|(X_{-l})_l\right],$$

$$\bar{\nu}_n^{(2,2)}(t) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{1}_{|\varepsilon_i|>\kappa_n} t \circ \hat{G}_n - \mathbb{E}\left[\varepsilon_i \mathbf{1}_{|\varepsilon_i|>\kappa_n} t \circ \hat{G}_n \,|(X_{-l})_l\right],$$

où

$$\kappa_n = c\sqrt{\frac{n}{\ln(n)}},$$

avec c une constante. Ainsi,

$$T_{2,b}^{m'} \leq \sum_{l\in\{(1),(2,1),(2,2)\}} T_{2,b,l}^{m'},$$

avec, pour $l \in \{(1),(2,1)\}$,

$$T_{2,b,l}^{m'} = D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \kappa' \frac{\ln(n)}{n}\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}(m')}\left(\bar{\nu}_n^l(t)\right)^2 - \frac{\kappa}{2\|\varphi_2'\|_\infty^2}\frac{D_{m'}}{n}\ln(n)\right)_+\right],$$

et

$$T_{2,b,(2,2)}^{m'} = D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \kappa' \frac{\ln(n)}{n}\mathbb{E}\left[\sup_{t\in\mathcal{S}(m')}\left(\bar{\nu}_n^{2,2}(t)\right)^2\right].$$

On va appliquer l'inégalité de Talagrand pour contrôler les déviations des deux premiers processus, qui sont bornés. Pour le processus $\bar{\nu}_n^{(1)}$, on notera $r_t(x) = f(x)t(\hat{G}_n(x))$. Calculons les quantités $M_1^{(1)}, H^{(1)}$ et $v^{(1)}$ telles que

$$\sup_{t\in\mathcal{S}(m')}\|r_t\|_\infty \leq M_1^{(1)},\ \mathbb{E}\left[\sup_{t\in\mathcal{S}(m')}\left|\bar{\nu}_n^{(1)}(t)\right| |(X_{-l})_l\right] \leq H^{(1)},$$
$$\text{et} \sup_{t\in\mathcal{S}(m')}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}\left(r_t(X_i)\,|(X_{-l})_l\,\right) \leq v^{(1)}.$$

Notons que dès que l'on considèrera $t \in \mathcal{S}(m')$, on écrira $t = \sum_{j=1}^{D_{m'}} b_j\varphi_j$, avec $\sum_j b_j^2 = 1$.

- Cherchons la constante $M_1^{(1)}$. Soit $t \in \mathcal{S}(m')$. On utilise la connexion de normes :

$$|r_t(x)| \quad\leq\quad \|t\|_\infty\|f\|_\infty \leq \phi_0\sqrt{D_m}\|t\|\|f\|_\infty \leq \phi_0\sqrt{D_m}\|f\|_\infty := M_1^{(1)}$$

- Cherchons la constante $H^{(1)}$. Par l'inégalité de Cauchy-Schwarz, il suffit de majorer la quantité

$$\mathbb{E}\left[\sup_{t\in\mathcal{S}_{m'}}\left(\bar{\nu}_n^{(1)}(t)\right)^2 |(X_{-l})_l\right].$$

En utilisant la décomposition d'une fonction t de $\mathcal{S}(m')$ dans la base $(\varphi_j)_j$ de l'espace $S_{m'}$, on obtient,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in\mathcal{S}_{m'}}\left(\bar{\nu}_n^{(1)}(t)\right)^2|(X_{-l})_l\right] &\leq \sum_{j=1}^{D_{m'}}\mathbb{E}\left[\left(\bar{\nu}_n^{(1)}(\varphi_j)\right)^2|(X_{-l})_l\right],\\
&= \frac{1}{n}\sum_{j=1}^{D_{m'}}\mathrm{Var}\left(f(X_i)\varphi_j(\hat{G}_n(X_i))\,|(X_{-l})_l\right),\\
&\leq \frac{1}{n}\sum_{j=1}^{D_{m'}}\mathbb{E}\left[f^2(X_i)\varphi_j^2(\hat{G}_n(X_i))\,|(X_{-l})_l\right],\\
&= \frac{1}{n}\left\|\sum_{j=1}^{D_{m'}}\varphi_j\right\|_\infty^2\mathbb{E}\left[f^2(X_1)\right],\\
&\leq \phi_0^2\mathbb{E}\left[f^2(X_1)\right]\frac{D_{m'}}{n} := \left(H^{(1)}\right)^2.
\end{aligned}$$

– Il reste à déterminer la constante $v^{(1)}$: soit $t\in\mathcal{S}(m')$,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}\left(r_t\left(X_i\right)|(X_{-l})_l\right) &= \mathrm{Var}\left(r_t\left(X_1\right)|(X_{-l})_l\right),\\
&\leq \mathbb{E}\left[f^2(X_i)t^2(\hat{G}_n(X_i))\,|(X_{-l})_l\right],\\
&= \mathbb{E}\left[\left\{\sum_{j=1}^{D_{m'}}b_j\varphi_j\left(\hat{G}_n(X_1)\right)f^2(X_1)\right\}^2|(X_{-l})_l\right],\\
&\leq \mathbb{E}\left[\sum_{j=1}^{D_{m'}}\left\{\varphi_j\left(\hat{G}_n(X_1)\right)f(X_1)\right\}^2|(X_{-l})_l\right],\\
&\leq n\left(H^{(1)}\right)^2 := v^{(1)}
\end{aligned}$$

On en déduit que l'inégalité de Talagrand s'applique : on obtient ainsi, pour $\delta=\kappa^{bis,1}\ln(n)$ $(\kappa^{bis,1}=8)$,

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}_{m'}}\left(\bar{\nu}_n^{(1)}(t)\right)^2-2(1+2\delta)\left(H^{(1)}\right)^2\right)_+|(X_{-l})_l\right]$$

$$\leq\frac{4}{K_1}\left\{\frac{v^{(1)}}{n}\exp\left(-K_1\delta\frac{n\left(H^{(1)}\right)^2}{v^{(1)}}\right)+\frac{49}{K_1C^2(\delta)}\frac{\left(M_1^{(1)}\right)^2}{n^2}\exp\left(\frac{-\sqrt{2}K_1C(\delta)\sqrt{\delta}}{7}\frac{nH^{(1)}}{M_1^{(1)}}\right)\right\},$$

où $C(\delta) = (\sqrt{1+\delta}-1)\wedge 1$ et $K_1 = 1/6$. Les ordres de grandeur des quantités clés sont les suivants :

$$\frac{v^{(1)}}{n} = \phi_0^2 \mathbb{E}\left[f^2(X_1)\right]\frac{D_{m'}}{n}, \quad \frac{n\left(H^{(1)}\right)^2}{v^{(1)}} = 1,$$
$$\frac{nH^{(1)}}{M_1^{(1)}} = \sqrt{n}\frac{\mathbb{E}\left[f^2(X_1)\right]}{\|f\|_\infty}, \qquad \frac{\left(M_1^{(1)}\right)^2}{n^2} = \phi_0^2\|f\|_\infty \frac{D_{m'}}{n^2}.$$

Ainsi, $\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}_{m'}}\left(\bar\nu_n^{(1)}(t)\right)^2 - 2(1+2\delta)\left(H^{(1)}\right)^2\right)_+ |(X_{-l})_l\right]$

$$\leq C\left\{\frac{D_{m'}}{n}\exp\left(-K_1\delta\right) + \frac{1}{C^2(\delta)}\frac{D_{m'}}{n^2}\exp\left(-CC(\delta)\sqrt{\delta}\sqrt{n}\right)\right\}.$$

Il vient, en sommant sur m', en notant que $C(\delta) = 1$ ici, et $D_{m'} \leq n^{1/3}$,

$$\begin{aligned}
C\frac{\ln(n)}{n}\sum_{m'\in\mathcal{M}_n} D_{m'}^3 \mathbb{E}&\left[\left(\sup_{t\in\mathcal{S}_{m'}}\left(\bar\nu_n^{(1)}(t)\right)^2 - 2(1+2\delta)\left(H^{(1)}\right)^2\right)_+ |(X_{-l})_l\right]\\
&\leq C\left\{\sum_{m'\in\mathcal{M}_n}\frac{D_{m'}^4}{n}n^{-K_1\kappa^{bis,1}} + \frac{1}{n^2}\sum_{m'\in\mathcal{M}_n} D_{m'}\exp(-C\sqrt{\kappa^{bis,1}}\sqrt{n}\sqrt{\ln(n)})\right\},\\
&\leq C\ln(n)\left(n^{(1/3)-\kappa^{bis,1}K_1} + n^{-5/3}\exp\left(-C\sqrt{\kappa^{bis,1}}\sqrt{n}\right)\right),\\
&\leq C\ln(n)\left(n^{(1/3)-\kappa^{bis,1}K_1} + n^{-1}\right),\\
&\leq C\frac{\ln(n)}{n},
\end{aligned}$$

si $n^{-\kappa^{bis,1}K_1+1/3} \leq n^{-1}$, ce qui est le cas puisque $\kappa^{bis,1} = 8$. On note aussi que

$$\begin{aligned}
2(1+2\delta)\left(H^{(1)}\right)^2 &\leq 6\delta\left(H^{(1)}\right)^2 = 6\kappa^{bis,1}\phi_0^2\mathbb{E}\left[f^2(X_1)\right]\ln(n)\frac{D_{m'}}{n},\\
&\leq 6\kappa^{bis,1}\phi_0^2\mathbb{E}\left[Y_1^2\right]\ln(n)\frac{D_{m'}}{n},\\
&\leq \frac{\kappa}{2\|\varphi_2'\|_\infty^2}\frac{D_{m'}}{n}\ln(n)
\end{aligned}$$

dès que $\kappa \geq 6\kappa^{bis,1}\phi_0^2\mathbb{E}[Y_1^2]2\|\varphi_2'\|_\infty^2$, ce qui est le cas. On a donc aussi, en prenant l'espérance,

$$\sum_{m'\in\mathcal{M}n} T_{2,b,(1)}^{m'} \leq \frac{C\ln(n)}{n}.$$

Pour le terme $T_{2,b,(2,1)}$, on applique à nouveau l'inégalité de Talagrand, pour le processus $\bar\nu_n^{(2,1)}$ cette fois. On notera $r_t(x,\varepsilon) = \varepsilon\mathbf{1}_{|\varepsilon|\leq\kappa_n}t(\hat G_n(x))$. Calculons les quantités $M_1^{(2)}, H^{(2)}$ et $v^{(2)}$ intervenant dans l'inégalité.

– Pour $t\in\mathcal{S}(m')$, on a :

$$|r_t(x,\varepsilon)| \quad\leq\quad \kappa_n\phi_0\sqrt{D_m} := M_1^{(2)}$$

– On majore ensuite

$$\mathbb{E}\left[\sup_{t\in\mathcal{S}_{m'}}\left(\bar\nu_n^{(2,1)}(t)\right)^2|(X_{-l})_l\right].$$

En utilisant la décomposition d'une fonction t de $\mathcal{S}(m')$ dans la base $(\varphi_j)_j$ de l'espace $S_{m'}$, on obtient,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in\mathcal{S}_{m'}}\left(\bar{\nu}_n^{(2,1)}(t)\right)^2|(X_{-l})_l\right] &\leq \sum_{j=1}^{D_{m'}}\mathbb{E}\left[\left(\bar{\nu}_n^{(2,1)}(\varphi_j)\right)^2|(X_{-l})_l\right],\\
&\leq \frac{1}{n}\sum_{j=1}^{D_{m'}}\mathbb{E}\left[\varepsilon_1^2\varphi_j^2(\hat{G}_n(X_i))\,|(X_{-l})_l\right],\\
&= \frac{1}{n}\left\|\sum_{j=1}^{D_{m'}}\varphi_j\right\|^2_{\infty,[0;1]}\mathbb{E}\left[f^2(X_1)\right],\\
&\leq \phi_0^2\mathbb{E}\left[\varepsilon_1^2\right]\frac{D_{m'}}{n}=\phi_0^2\sigma^2\frac{D_{m'}}{n}:=\left(H^{(2)}\right)^2.
\end{aligned}$$

– Pour $t\in\mathcal{S}(m')$,

$$\begin{aligned}
\mathrm{Var}\left(r_t\left(X_1\right)|(X_{-l})_l\right) &\leq \mathbb{E}\left[\varepsilon_i^2t^2(\hat{G}_n(X_i))\,|(X_{-l})_l\right],\\
&\leq \mathbb{E}\left[\sum_{j=1}^{D_{m'}}\left\{\varphi_j\left(\hat{G}_n(X_1)\right)\varepsilon_1\right\}^2|(X_{-l})_l\right],\\
&\leq n\left(H^{(2)}\right)^2:=v^{(2)}
\end{aligned}$$

On en déduit que l'inégalité de Talagrand s'applique : on obtient ainsi, pour $\delta=\kappa^{bis,2}\ln(n)$ ($\kappa^{bis,2}=8$), et avec les quantités clés suivantes,

$$\begin{aligned}
&\frac{v^{(2)}}{n}=\phi_0^2\sigma^2\frac{D_{m'}}{n}, \quad \frac{n\left(H^{(2)}\right)^2}{v^{(2)}}=1,\\
&\frac{nH^{(2)}}{M_1^{(2)}}=\frac{\sqrt{n}}{\kappa_n}\sigma, \qquad \frac{\left(M_1^{(2)}\right)^2}{n^2}=\phi_0^2\frac{\kappa_nD_{m'}}{n^2},
\end{aligned}$$

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{S}_{m'}}\left(\bar{\nu}_n^{(2,1)}(t)\right)^2-2(1+2\delta)\left(H^{(2)}\right)^2\right)_+|(X_{-l})_l\right]$$

$$\leq C\left\{\frac{D_{m'}}{n}\exp\left(-K_1\delta\right)+\frac{1}{C^2(\delta)}\frac{\kappa_nD_{m'}}{n^2}\exp\left(-CC(\delta)\sqrt{\delta}\frac{\sqrt{n}}{\kappa_n}\right)\right\}.$$

Il vient, en sommant sur m', en notant que $C(\delta)=1$ ici, et $D_{m'} \le n^{1/3}$,

$$C\frac{\ln(n)}{n}\sum_{m'\in\mathcal{M}_n} D_{m'}^3 \mathbb{E}\left[\left(\sup_{t\in\mathcal{S}_{m'}}\left(\bar\nu_n^{2,1}(t)\right)^2 - 2(1+2\delta)\left(H^{(1)}\right)^2\right)_+ |(X_{-l})_l\right]$$

$$\begin{aligned}
&\le C\left\{\sum_{m'\in\mathcal{M}_n}\frac{D_{m'}^4}{n}n^{-K_1\kappa^{bis,2}} + \frac{\kappa_n}{n^2}\sum_{m'\in\mathcal{M}_n} D_{m'}\exp(-C\sqrt{\kappa^{bis,2}}\frac{\sqrt{n}}{\kappa_n}\sqrt{\ln(n)})\right\},\\
&\le C\ln(n)\left(n^{(1/3)-\kappa^{bis,1}K_1} + \frac{n^{-7/6}}{\sqrt{\ln(n)}}n^{-C\sqrt{\kappa^{bis,2}}}\right),\\
&\le C\ln(n)\left(n^{(1/3)-\kappa^{bis,1}K_1} + n^{-1}\right),\\
&\le C\frac{\ln(n)}{n},
\end{aligned}$$

si $n^{-\kappa^{bis,2}K_1+1/3} \le n^{-1}$, ce qui est le cas puisque $\kappa^{bis,2}=8$. On note aussi que

$$\begin{aligned}
2(1+2\delta)\left(H^{(2)}\right)^2 &\le 6\delta\left(H^{(1)}\right)^2 = 6\kappa^{bis,2}\phi_0^2\sigma^2\ln(n)\frac{D_{m'}}{n},\\
&\le 6\kappa^{bis,2}\phi_0^2\mathbb{E}\left[Y_1^2\right]\ln(n)\frac{D_{m'}}{n},\\
&\le \frac{\kappa}{2\|\varphi_2'\|_\infty^2}\frac{D_{m'}}{n}\ln(n)
\end{aligned}$$

dès que $\kappa \ge 6\kappa^{bis,1}\phi_0^2\mathbb{E}[Y_1^2]2\|\varphi_2'\|_\infty^2$, ce qui est le cas. On a donc aussi, en prenant l'espérance,

$$\sum_{m'\in\mathcal{M}n} T_{2,b,(2,1)}^{m'} \le \frac{C\ln(n)}{n}.$$

Enfin, pour le troisième terme, on se contente de majorations plus grossières :

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in\mathcal{S}(m')}\left(\bar\nu_n^{2,2}(t)\right)^2|(X_{-l})_l\right] &\le \sum_{j=1}^{D_{m'}}\mathbb{E}\left[\left(\bar\nu_n^{2,2}(\varphi_j)\right)^2|(X_{-l})_l\right],\\
&\le \frac{1}{n}\sum_{j=1}^{D_{m'}}\mathbb{E}\left[\varepsilon_1^2\mathbf{1}_{|\varepsilon_1|>\kappa_n}\varphi_j^2\circ\hat G_n(X_1)\,|(X_{-l})_l\right],\\
&\le \phi_0^2\frac{D_{m'}}{n}\mathbb{E}\left[\varepsilon_1^2\mathbf{1}_{|\varepsilon_1|>\kappa_n}\right],\\
&\le \phi_0^2\kappa_n^{-p}\frac{D_{m'}}{n}\mathbb{E}\left[|\varepsilon_1|^{2+p}\right],
\end{aligned}$$

pour tout $p>0$. Par conséquent,

$$\begin{aligned}
\sum_{m'\in\mathcal{M}_n} T_{2,b,(2,2)}^{m'} &\le C\frac{\ln(n)}{n}\sum_{m'\in\mathcal{M}_n}\frac{D_{m'}^4}{n}\kappa_n^{-p},\\
&\le Cn^{1/3-p/2}\ln(n)^{1-p/2} \le C\ln(n)n^{-1},
\end{aligned}$$

dès que $1/3-p/2\le -1$, ie. $p\ge 8/3$, ce qui est possible, puisque l'on a supposé $\varepsilon_1\in L^{2+8/3}(\mathbb{P})$. Finalement,

$$\sum_{m'\in\mathcal{M}_n} T_{2,b}^{m'} \le C\frac{\ln(n)}{n},$$

ce qui termine la preuve du lemme.

□

9.7.2. *Preuve du Lemme 27.*
Commençons par noter que $V_3(p_{m'}) \leq V_3(m')$. Par conséquent,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{p_{m'},b} - V_3(m')\right)_+\right] \leq \mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{p_{m'},b} - V_3(p_{m'})\right)_+\right].$$

On note $p = p_{m'}$ pour simplifier les notations dans la suite. On commence par effectuer la même décomposition qu'en non-adaptatif (cf. Lemme 23) :

$$T_3^{p,b} \leq 6T_{3,1,1}^p + 6T_{3,1,2}^p + 3T_{3,2}^p + 3T_{3,3}^p.$$

Ainsi,

$$\begin{aligned}\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_3^{p,b} - V_3(p)\right)_+\right] \leq \mathbb{E}&\left[\max_{m'\in\mathcal{M}_n}\left(6T_{3,1,1}^p - V_3(p)/3\right)_+\right]\\ &+\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} 6T_{3,1,2}^p\right]\\ &+\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,2}^p - V_3(p)/3\right)_+\right]\\ &+\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,3}^p - V_3(p)/3\right)_+\right].\end{aligned}$$

L'un des termes n'a pas été recentrer, il est directement négligeable : la ligne de calcul (22) prouve que, quel que soit $m' \in \mathcal{M}_n$,

$$T_{3,1,2}^p \leq T_{3,1,2}^{m_{\max}},$$

et les calculs qui suivent justifient donc que :

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} 6T_{3,1,2}^p\right] \leq \frac{C}{n}.$$

Pour traiter les termes recentrés, on utilise les lemmes suivants :

Lemme 30. *Sous les hypothèses du Théorème 5, il existe une constante numérique C telle que,*

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(6T_{3,1,1}^p - V_{3,1,1}(p)\right)_+\right] \leq \frac{C}{n},$$

avec

$$V_{3,1,1}(p) = 6 \times 2(1+2\delta)\mathbb{E}\left[f^2(X_1)\right]\phi_0^2\frac{D_p}{n}.$$

De plus,

$$V_{3,1,1}(p) \leq 12(1+2\delta)\mathbb{E}\left[Y_1^2\right]\phi_0^2\frac{D_p}{n} := V_{3,1,1}^{bis}(p).$$

L'inégalité du lemme est encore valable pour $V_{3,1,1}$ remplacé par $V_{3,1,1}^{bis}$.

Lemme 31. *Sous les hypothèses du Théorème 5, il existe une constante numérique C, dépendant de $\|h\|^2$ telle que*

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,2}^p - V_{3,2}(p)\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec pour $\kappa = 50\pi^4/3\|h\|^2$,

$$V_{3,2}(p) = \kappa \frac{D_p^4 \ln^2(n)}{n^2}.$$

On note que

$$V_{3,2}(p) \leq 50\pi^4/3\mathbb{E}\left[Y_1^2\right] \frac{D_p^4 \ln^3(n)}{n^2}.$$

Si de plus, $D_p \leq n^{1/3}/\ln(n)$, on a

$$V_{3,2}(p) \leq 50\pi^4/3\mathbb{E}\left[Y_1^2\right] \frac{D_p}{n} := V_{3,2}^{bis}(p).$$

L'inégalité du lemme est encore valable pour $V_{3,2}$ remplacé par $V_{3,2}^{bis}$.

Lemme 32. *Sous les hypothèses du Théorème 5, il existe une constante numérique C, dépendant de $\|h\|^2$ telle que*

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,3}^p - V_{3,3}(p)\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec pour $\kappa = (13^3 * 2/27)\|h\|^2\|\varphi_2^{(3)}\|_\infty^2$,

$$V_{3,3}(p) = \kappa \frac{D_p^7 \ln^3(n)}{n^3}.$$

On note que

$$V_{3,3}(p) \leq (13^3 \times 2/27)\|\varphi_2^{(3)}\|_\infty^2\mathbb{E}\left[Y_1^2\right] \frac{D_p^7 \ln^4(n)}{n^3}.$$

Si de plus, $D_p \leq n^{1/3}/\ln^{2/3}(n)$, on a

$$V_{3,3}(p) \leq (13^3 \times 2/27)\|\varphi_2^{(3)}\|_\infty^2\mathbb{E}\left[Y_1^2\right] \frac{D_p}{n} := V_{3,3}^{bis}(p).$$

L'inégalité du lemme est encore valable pour $V_{3,3}$ remplacé par $V_{3,3}^{bis}$.
Pour conclure, on note que le choix d'une constante c_3 assez grande assure que $V_3 \geq 3V_{3,1,1}^{bis}$, $V_3 \geq 3V_{3,2}^{bis}$, et $V_3 \geq 3V_{3,3}^{bis}$. Ainsi, les inégalités des Lemmes 30, 31 et 32 sont encore valables en substituant V_3 aux quantités $V_{3,i}^{bis}$, $i \in \{(1,1), 2, 3\}$.

□

9.7.3. *Preuve du Lemme 28.*
On obtient déjà la décomposition, comme en non adaptatif (cf. Lemme 24)

$$T_4^{m'} \leq 4T_{4,1,1}^{m'} + 4T_{4,1,2}^{m'} + 2T_{4,2,1}^{m'} + 2T_{4,2,2}^{m'} + 2T_{4,2,3}^{m'},$$

et donc aussi,

$$\begin{aligned}
\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_4^{m'} - V_4(m')\right)_+\right] \quad &\leq \quad \mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(4T_{4,1,1}^{m'} - V_4(m')/3\right)_+\right] \\
&\quad +\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(4T_{4,1,2}^{m'} - V_4(m')/3\right)_+\right] \\
&\quad +\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(2T_{4,2,3}^{m'} - V_4(m')/3\right)_+\right] \\
&\quad +\mathbb{E}\left[\max_{m'\in\mathcal{M}_n} 2T_{4,2,1}^{m'}\right] + \mathbb{E}\left[\max_{m'\in\mathcal{M}_n} 2T_{4,2,2}^{m'}\right].
\end{aligned}$$

Deux des termes n'ont pas été recentrés, car ils sont négligeables devant $\ln(n)/n$ sans contrainte supplémentaire que $D^{(1)}_{m_{1,\max}} \leq n^{1/3}$. On a obtenu en effet, à la ligne (27)

$$T^{m'}_{4,2,1} \quad \leq \quad \left\| id - \hat{U}_n \right\|^2_\infty \left\| h' \right\|^2,$$

et donc, le majorant étant indépendant de m',

$$\mathbb{E}\left[\max_{m' \in \mathcal{M}_n} T^{m'}_{4,2,1}\right] \leq C_2 \left\| h' \right\|^2 \frac{1}{n} = \frac{C}{n}.$$

De même, on obtient, à l'aide des lignes (28) et (29)

$$\begin{aligned} T^{m'}_{4,2,2} \quad &\leq \quad (1/4) \int_0^1 (u - \hat{U}_n(u))^4 \|\varphi_2''\|^2_\infty \left(\sum_{j=1}^{D_m} a_j^2 \mu_j^2 \right) \sum_{j=1}^{D_m} \mu_j^2 du, \\ &\leq \quad (1/4) \int_0^1 (u - \hat{U}_n(u))^4 \|\varphi_2''\|^2_\infty \left(\sum_{j=1}^{D_{m_{\max}}} a_j^2 \mu_j^2 \right) \sum_{j=1}^{D_{m_{\max}}} \mu_j^2 du, \end{aligned}$$

et donc aussi

$$\mathbb{E}\left[\max_{m' \in \mathcal{M}_n} T^{m'}_{4,2,2}\right] \leq (1/4) \left\| \varphi_2'' \right\|^2_\infty \frac{L^2}{\pi^2} C_4 \frac{D^3_{m_{\max}}}{n^2} \leq \frac{C}{n},$$

dès que $D_{m_{\max}} \leq n^{1/3}$. Il reste à majorer les termes recentrés. On utilise pour cela les lemmes suivants. Tout d'abord, concernant le terme où $T^{m'}_{4,1,1}$ intervient :

Lemme 33. *Sous les hypothèses du Théorème 5 :*

$$\mathbb{E}\left[\max_{m' \in \mathcal{M}_n} \left(T^{m'}_{4,1,1} - V_{4,1,1}(m') \right)_+ \right] \leq C \frac{\ln(n)}{n},$$

où l'on note

$$V_{4,1,1}(m') = \kappa \frac{D^4_{m'} \ln(n)}{n^2},$$

avec $\kappa = (10/3)\phi_0^2 \|\varphi_2'\|^2_\infty \mathbb{E}[Y_1^2]$.

Sous l'hypothèse $D_{m'} \leq (n/\ln(n))^{1/3}$, on a

$$V_{4,1,1}(m') \leq \kappa \frac{D_{m'}}{n} := V^{bis}_{4,1,1}(m').$$

L'inégalité du lemme est encore valable pour $V_{4,1,1}$ remplacé par $V^{bis}_{4,1,1}$.

Concernant le terme,

$$T^{m'}_{4,1,2} = T^{m',b}_3 \int_{[0;1]} \sum_{j=1}^{D_{m'}} \left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u) \right)^2 du,$$

on commence par remplacer $T^{m',b}_3$ par son majorant détaillé :

$$T^{m',b}_3 \leq 6T^{m'}_{3,1,1} + 6T^{m'}_{3,1,2} + 3T^{m'}_{3,2} + 3T^{m'}_{3,3},$$

et majorer :

$$\int_{[0;1]} \sum_{j=1}^{D_{m'}} \left(\varphi_j(u) - \varphi_j \circ \hat{U}_n(u) \right)^2 du \leq D^3_{m'} \left\| \varphi_2' \right\|^2_\infty \left\| \hat{U}_n - id \right\|^2_\infty.$$

Cela donne donc,

$$T_{4,1,2}^{m'} \leq \sum_{l=1}^{4} T_{4,1,2,l}^{m'},$$

avec

$$\begin{aligned}
T_{4,1,2,1}^{m'} &= 6T_{3,1,1}^{m'} D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \left\|\hat{U}_n - id\right\|_\infty^2, \\
T_{4,1,2,2}^{m'} &= 6T_{3,1,2}^{m'} D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \left\|\hat{U}_n - id\right\|_\infty^2, \\
T_{4,1,2,3}^{m'} &= 3T_{3,2}^{m'} D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \left\|\hat{U}_n - id\right\|_\infty^2, \\
T_{4,1,2,4}^{m'} &= 3T_{3,3}^{m'} D_{m'}^3 \left\|\varphi_2'\right\|_\infty^2 \left\|\hat{U}_n - id\right\|_\infty^2.
\end{aligned}$$

Puis on utilise :

Lemme 34. *Sous les hypothèses du Théorème 5, pour chaque $l = 1, \dots, 4$, il existe une constante numérique C, dépendant de $\|h\|^2$ telle que*

$$\mathbb{E}\left[\max_{m' \in \mathcal{M}_n} \left(T_{4,1,2,l}^{m'} - V_{4,1,2,l}(m')\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec

$$\begin{aligned}
V_{4,1,2,1}(m') &= \kappa_0\kappa_1 \frac{D_{m'}^4 \ln(n)}{n^2}, \\
V_{4,1,2,2}(m') &= \kappa_2 \frac{D_{m'}^3 \ln^2(n)}{n^2}, \\
V_{4,1,2,3}(m') &= \kappa_3 \frac{D_{m'}^7 \ln^3(n)}{n^3}, \\
V_{4,1,2,4}(m') &= \kappa_4 \frac{D_{m'}^{10} \ln^4(n)}{n^4},
\end{aligned}$$

et les constantes

$$\begin{aligned}
&\kappa_0 = 10/3,\ \kappa_1 = \|\varphi_2'\|_\infty^2 \times 12(1+2\delta)\mathbb{E}\left[f^2(X_1)\right]\phi_0^2, \\
&\kappa_2 = 18 \times 6\|h'\|^2\|\varphi_2'\|_\infty^2, \\
&\kappa_3 = 13^3/27 \times 4 \times (3\pi^4/4)\|h\|^2\|\varphi_2'\|_\infty^2, \\
&\kappa_4 = 16^4/81 \times 8 \times (1/2)\|h\|^2\|\varphi_2^{(3)}\|_\infty^2\|\varphi_2'\|_\infty^2.
\end{aligned}$$

On se débarrasse des constantes dépendant de h, inconnue, et on majore les $V_{4,1,2,l}$, sous certaines contraintes sur les dimensions, par des quantités de l'ordre D_m/n. Pour cela, on note que $\mathbb{E}[f^2(X_1)] = \|h\|^2 \leq \mathbb{E}[Y_1^2]$. Ainsi,

$$V_{4,1,2,1}(m') \quad \leq \quad 40\phi_0^2\|\varphi_2'\|_\infty^2(1+2\delta)\mathbb{E}\left[Y_1^2\right]\frac{D_{m'}}{n} = V_{4,1,2,1}^{bis}(m'),$$

pour $D_{m'} \leq (n/\ln(n))^{1/3}$.
De même

$$V_{4,1,2,3}(m') \quad \leq \quad 13^3/27 \times (3\pi^4)\|\varphi_2'\|_\infty^2\mathbb{E}\left[Y_1^2\right]\frac{D_{m'}}{n} := V_{4,1,2,3}^{bis}(m'),$$

pour $D_{m'} \leq n^{1/3}/\ln(n)^{1/2}$, et aussi

$$V_{4,1,2,4}(m') \quad \leq \quad 16^4/81 \times 4\|\varphi_2^{(3)}\|_\infty^2\|\varphi_2'\|_\infty^2\mathbb{E}\left[Y_1^2\right]\frac{D_{m'}}{n},$$

pour $D_{m'} \leq n^{1/3}/\ln^{4/9}(n)$.
Enfin, pour le terme $V_{4,1,2,1}$, on se débarrasse de la norme de la dérivée de h au prix d'une condition sur le nombre d'observations n : pour $n \geq n_1 = \exp(\|h'\|^2)$,

$$V_{4,1,2,2}(m') \leq 6\|\varphi_2'\|_\infty^2 \leq \frac{D_{m'}^3 \ln^3(n)}{n^2} \leq 6\|\varphi_2'\|_\infty^2 \leq \frac{D_{m'}}{n} := V_{4,1,2,2}^{bis}(m'),$$

pour $D_{m'} \leq n^{1/3}/\ln(n)$.
Les inégalités du lemme sont encore valables en substituant les quantités $V_{4,1,2,l}^{bis}$ aux quantités $V_{4,1,2,l}$.

Enfin, pour $T_{4,2,3}^{m'}$:

Lemme 35. *Sous les hypothèses du Théorème 5, il existe une constante numérique C, dépendant de $\|h\|^2$ telle que*

$$\mathbb{E}\left[\max_{m' \in \mathcal{M}_n} \left(T_{4,2,3}(m') - V_{4,2,3}(m')\right)_+\right] \leq C\frac{\ln(n)}{n},$$

avec pour $\kappa = (8/3)^{3/2} 2^{1/2} \|\varphi_2'\|_\infty \|\varphi_2''\|_\infty L^2/\pi^2$,

$$V_{4,2,3}(m') = \kappa \frac{D_{m'}^2 \ln^{3/2}(n)}{n^{3/2}}.$$

Si de plus, $D_{m'} \leq n^{1/2}/\ln(n)^{3/2}$, on a

$$V_{4,2,3}(m') \leq \kappa \frac{D_{m'}}{n} := V_{4,2,3}^{bis}(m').$$

L'inégalité du lemme est encore valable pour $V_{4,2,3}$ remplacé par $V_{4,2,3}^{bis}$. Pour conclure, on note que le choix d'une constante c_4 assez grande assure que $V_4 \geq 12V_{4,1,1}^{bis}$, $V_4 \geq 48V_{4,1,2,l}^{bis}$ $(l = 1, \ldots, 4)$, et $V_4 \geq 6V_{4,2,3}^{bis}$. Ainsi, les inégalités des Lemmes 33, 34 et 35 sont encore valables en substituant V_4 aux quantités $V_{4,i}^{bis}$, $i \in \{(1,1), (1,2,l), l = 1, \ldots, 4, (2,3)\}$.

□

9.7.4. *Preuve du Lemme 30.*
On constate qu'on peut écrire,

$$T_{3,1,1}^p = \sum_{j=1}^{D_p} \left\{\nu_n^{(1)}(\varphi_j)\right\}^2,$$

où $\nu_n^{(1)}$ est le processus empirique suivant : pour $s \in L^2([0;1])$

$$\nu_n^{(1)}(s) = \frac{1}{n}\sum_{i=1}^n f(X_i) s \circ G(X_i) - \mathbb{E}\left[f(X_i) s \circ G(X_i)\right],$$

en utilisant les définitions $U_i = G(X_{-i})$, $f = h \circ G$. Or, on a l'égalité

$$\sum_{j=1}^{D_p} \left\{\nu_n^{(1)}(\varphi_j)\right\}^2 = \sup_{s \in S_p,\ \|s\|=1} \left(\nu_n^{(1)}(s)\right)^2. \tag{36}$$

On applique l'inégalité de Talagrand dans sa version intégrée au processus $\nu_n^{(1)}$, comme dans la preuve du Lemme Clé :

$$\mathbb{E}\left[\left(\sup_{s\in\mathcal{S}_p}\left(\nu_n^{(1)}(t)\right)^2-2(1+2\delta)\left(H^{(1)}\right)^2\right)_+\right]$$
$$\leq\frac{4}{K_1}\left\{\frac{v^{(1)}}{n}\exp\left(-K_1\delta\frac{n\left(H^{(1)}\right)^2}{v^{(1)}}\right)+\frac{49}{K_1C^2(\delta)}\frac{\left(M_1^{(1)}\right)^2}{n^2}\exp\left(\frac{-\sqrt{2}K_1C(\delta)\sqrt{\delta}}{7}\frac{nH^{(1)}}{M_1^{(1)}}\right)\right\},$$

où $C(\delta)=(\sqrt{1+\delta}-1)\wedge 1$ et $K_1=1/6$, et avec les quantités :

$$M_1^{(1)}=\phi_0\|f\|_\infty\sqrt{D_p},\ \left(H^{(1)}\right)^2=\frac{D_p}{n}\mathbb{E}\left[f^2(X_1)\right]\phi_0^2,\ v^{(1)}=\|f\|_\infty^2.$$

Les ordres de grandeur des quantités clés sont les suivants :

$$\frac{v^{(1)}}{n}=\frac{\|f\|_\infty^2}{n},\qquad \frac{n(H^{(1)})^2}{v}=\frac{D_p\left[f^2(X_1)\right]\phi_0^2}{\|f\|_\infty^2},$$
$$\frac{nH^{(1)}}{M_1^{(1)}}=\sqrt{n}\frac{\sqrt{\mathbb{E}[f^2(X_1)]}}{\|f\|_\infty},\qquad \frac{(M_1^{(1)})^2}{n^2}=\frac{\phi_0^2\|f\|_\infty^2D_p}{n^2}.$$

On conclut ensuite classiquement que

$$\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(\sup_{\substack{s\in S_p\\ \|s\|=1}}\left(\nu_n^{(1)}(s)\right)^2-2(1+2\delta)\left(H^{(1)}\right)^2\right)_+\right]\leq\frac{C}{n}.$$

□

9.7.5. *Preuve du Lemme 31.*
On commence comme toujours par

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,2}^p-V_{3,2}(p)\right)_+\right]\leq\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(3T_{3,2}^p-V_{3,2}(p)\right)_+\right],$$

et on raisonne pour p fixé. On a montré en non adaptatif que

$$T_{3,2}^p\leq(\pi^4/4)\|h\|^2D_p^4\left\|\hat{U}_n-id\right\|_\infty^4.$$

Par conséquent,

$$\begin{aligned}&\mathbb{E}\left[\left(3T_{3,2}^p-V_{3,2}(p)\right)_+\right]\\&\quad\leq\ (3\pi^4/4)\|h\|^2D_p^4\mathbb{E}\left[\left(\left\|\hat{U}_n-id\right\|_\infty^4-\frac{\kappa}{(3\pi^4/4)\|h\|^2}\frac{\ln^2(n)}{n^2}\right)_+\right],\\&\quad\leq\ CD_p^4n^{-\kappa_b^{1/2}2^{-1/2}},\end{aligned}$$

en utilisant le Corollaire 15 avec $p=4$, et $\kappa_b=\frac{\kappa}{(3\pi^4/4)\|h\|^2}$. Ainsi, si $D_p\leq n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T_{3,2}^p-V_{3,2}(p)\right)_+\right]\leq Cn\times n^{4/3}\times n^{-\kappa_b^{1/2}2^{-1/2}}.$$

Le choix de $\kappa=50\pi^4/3\|h\|^2$ entraîne $\kappa_b\geq 200/9$, et donc $7/3-\sqrt{\kappa_b/2}\leq -1$. Ce qui termine la preuve.

□

9.7.6. *Preuve du Lemme 32.*
On commence comme toujours par

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T^p_{3,3}-V_{3,3}(p)\right)_+\right]\le\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(3T^p_{3,3}-V_{3,3}(p)\right)_+\right],$$

et on raisonne pour p fixé. On a montré en non adaptatif que

$$T^p_{3,3}\le(1/6)\left\|\varphi_2^{(3)}\right\|^2_\infty\|h\|^2D_p^7\left\|\hat{U}_n-id\right\|^6_\infty.$$

Par conséquent,

$$\begin{aligned}&\mathbb{E}\left[\left(3T^p_{3,3}-V_{3,3}(p)\right)_+\right]\\ &\quad\le\ (1/2)\|h\|^2\left\|\varphi_2^{(3)}\right\|^2_\infty D_p^7\mathbb{E}\left[\left(\left\|\hat{U}_n-id\right\|^6_\infty-\frac{\kappa}{(1/2)\|h\|^2\|\varphi_2^{(3)}\|^2_\infty}\frac{\ln^3(n)}{n^3}\right)_+\right],\\ &\quad\le\ CD_p^7n^{-\kappa_b^{1/3}2^{-2/3}},\end{aligned}$$

en utilisant le Corollaire 15 avec $p=6$, et $\kappa_b=\frac{\kappa}{(1/2)\|h\|^2\|\varphi_2^{(3)}\|^2_{\infty,[0;1]}}$. Ainsi, si $D_p\le n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(3T^p_{3,3}-V_{3,3}(p)\right)_+\right]\le Cn\times n^{7/3}\times n^{-\kappa_b^{1/3}2^{-2/3}}.$$

Le choix de $\kappa=(13^3*2/27)\|h\|^2\|\varphi_2^{(3)}\|^2_\infty$ entraîne $\kappa_b\ge 13^3/27*4$, et donc $10/3-\kappa_b^{1/3}\times 2^{-2/3}\le -1$. Ce qui termine la preuve.

□

9.7.7. *Preuve du Lemme 33.*
On majore d'abord

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T^{m'}_{4,1,1}-V_{4,1,1}(m')\right)_+\right]\le\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(T^{m'}_{4,1,1}-V_{4,1,1}(m')\right)_+\right].$$

Puis, on utilise la majoration (25) :

$$\begin{aligned}&\mathbb{E}\left[\left(T^{m'}_{4,1,1}-\kappa\frac{D^4_{m'}\ln(n)}{n^2}\right)_+\right]\\ &\quad\le\ \mathbb{E}\left[\left(\phi_0^2\mathbb{E}\left[Y_1^2\right]\frac{D^4_{m'}}{n}\|\varphi_2'\|^2_\infty\left\|\hat{U}_n-id\right\|^2_\infty-\kappa\frac{D^4_{m'}\ln(n)}{n^2}\right)_+\right],\\ &\quad=\ \frac{D^4_{m'}}{n}\mathbb{E}\left[\left(\phi_0^2\mathbb{E}\left[Y_1^2\right]\|\varphi_2'\|^2_\infty\left\|\hat{U}_n-id\right\|^2_\infty-\kappa\frac{\ln(n)}{n}\right)_+\right].\end{aligned}$$

On majore ce terme à l'aide du Corollaire 15, avec $p=2$:

$$\mathbb{E}\left[\left(\phi_0^2\left\|\varphi_2'\right\|^2_\infty\mathbb{E}\left[Y_1^2\right]\left\|\hat{U}_n-id\right\|^2_\infty-\kappa\frac{\ln(n)}{n}\right)_+\right]\le Cn^{-\frac{\kappa}{\phi_0^2\|\varphi_2'\|^2_\infty\mathbb{E}[Y_1^2]}}.$$

On conclut comme dans la preuve du Lemme 26 : on a donc, pour $D_m\le n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T^{m'}_{4,1,1}-V_{4,1,1}(m')\right)_+\right]\le C\sum_{m'\in\mathcal{M}_n}D^4_{m'}n^{-\frac{\kappa}{\phi_0^2\|\varphi_2'\|^2_\infty\mathbb{E}[Y_1^2]}}\le Cn\times n^{4/3}n^{-\frac{\kappa}{\phi_0^2\|\varphi_2'\|^2_\infty\mathbb{E}[Y_1^2]}},$$

et le choix de $\kappa = (10/3)\phi_0^2\|\varphi_2'\|_\infty^2\mathbb{E}[Y_1^2]$ assure que $7/3 - \kappa/(\phi_0^2\|\varphi_2'\|_\infty^2\mathbb{E}[Y_1^2]) \leq -1$, et donc le résultat du lemme suit.

$\square$

9.7.8. *Preuve du lemme 34.*
On commence comme toujours par

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_{4,1,2,l}^{m'} - V_{4,1,2,l}(m')\right)_+\right] \leq \sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(T_{4,1,2,l}^{m'} - V_{4,1,2,l}(m')\right)_+\right],$$

pour $l = 1, \ldots, 4$. Commençons par les trois termes d'indices $l = 2, 3, 4$. On utilise des majorations non adaptatives pour chacun des termes : on a montré,

$$\begin{aligned}
T_{3,1,2}^m &\leq \left\|h'\right\|^2\left\|\hat{U}_n - id\right\|_\infty^2,\\
T_{3,2}^m &\leq \frac{\pi^4}{4}D_m^4\left\|h\right\|^2\left\|\hat{U}_n - id\right\|_\infty^4,\\
T_{3,3}^m &\leq \frac{1}{6}D_m^7\left\|\varphi_2^{(3)}\right\|_\infty^2\left\|h\right\|^2\left\|\hat{U}_n - id\right\|_\infty^6.
\end{aligned}$$

Ceci entraîne, dans le cas $l = 2$,

$$\begin{aligned}
&\mathbb{E}\left[\left(T_{4,1,2,2}^{m'} - V_{4,1,2,2}(m')\right)_+\right]\\
&\leq\ 6\left\|h'\right\|^2\left\|\varphi_2'\right\|_\infty^2 D_{m'}^3\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^4 - \frac{\kappa_2}{6\|h'\|^2\|\varphi_2'\|_\infty^2}\frac{\ln^2(n)}{n^2}\right)_+\right],\\
&\leq\ CD_m^3 n^{-\kappa_{b,2}^{1/2}2^{-1/2}},
\end{aligned}$$

en utilisant le Corollaire 15 avec $p = 4$, et $\kappa_{b,2} = \frac{\kappa_2}{6\|h'\|^2\varphi_2'\|_\infty^2}$. Ainsi, si $D_{m'} \leq n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_{4,1,2,2}^{m'} - V_{4,1,2,2}(m')\right)_+\right] \leq Cn \times n \times n^{-\kappa_{b,2}^{1/2}2^{-1/2}}.$$

Le choix de $\kappa_2 = 18\times 6\|h'\|^2\|\varphi_2'\|_\infty^2$ entraîne $\kappa_{b,2} \geq 18$, et donc $2 - \sqrt{\kappa_{b,2}/2} \leq -1$. Ce qui termine la majoration de ce terme. On répète la méthode pour $l = 3$:

$$\begin{aligned}
&\mathbb{E}\left[\left(T_{4,1,2,3}^{m'} - V_{4,1,2,3}(m')\right)_+\right]\\
&\leq\ \frac{3\pi^4}{4}\left\|h\right\|^2\left\|\varphi_2'\right\|_\infty^2 D_{m'}^7\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^6 - \frac{\kappa_3}{(3\pi^4/4)\|h\|^2\|\varphi_2'\|_\infty^2}\frac{\ln^3(n)}{n^3}\right)_+\right],\\
&\leq\ CD_{m'}^7 n^{-\kappa_{b,3}^{1/2}2^{-1/2}},
\end{aligned}$$

en utilisant le Corollaire 15 avec $p = 6$, et $\kappa_{b,3} = \frac{\kappa_3}{(3\pi^4/4)\|h\|^2\varphi_2'\|_{\infty,[0;1]}^2}$. Ainsi, si $D_{m'} \leq n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_{4,1,2,3}^{m'} - V_{4,1,2,3}(m')\right)_+\right] \leq Cn \times n^{7/3} \times n^{-\kappa_{b,2}^{1/3}2^{-2/3}}.$$

Le choix de $\kappa_3 = 13^3/27 \times 4 \times (3\pi^4/4)\|h\|^2\|\varphi_2'\|_\infty^2$ entraîne $\kappa_{b,3} \geq 13^3/27 \times 4$, et donc $10/3 - \kappa_{b,3}^{1/3} \times 2^{-2/3} \leq -1$. Ce qui termine la majoration de ce terme. Enfin, pour $l = 4$,

$$
\begin{aligned}
&\mathbb{E}\left[\left(T^{m'}_{4,1,2,4}-V_{4,1,2,4}(m')\right)_+\right]\\
&\qquad\leq \frac{1}{2}\|h\|^2\left\|\varphi_2^{(3)}\right\|_\infty^2\left\|\varphi_2'\right\|_\infty^2 D_{m'}^{10}\\
&\qquad\quad\times\mathbb{E}\left[\left(\left\|\hat{U}_n-id\right\|_\infty^8-\frac{\kappa_4}{(1/2)\|h\|^2\|\varphi_2^{(3)}\|_\infty^2\|\varphi_2'\|_\infty^2}\frac{\ln^4(n)}{n^4}\right)_+\right],\\
&\qquad\leq CD_{m'}n^{-\kappa_{b,4}^{1/4}2^{-3/4}},
\end{aligned}
$$

en utilisant le Corollaire 15 avec $p=8$, et $\kappa_{b,4}=\frac{\kappa_4}{(1/2)\|h\|^2\|\varphi_2^{(3)}\|_\infty^2\|\varphi_2'\|_\infty^2}$. Ainsi, si $D_{m'}\leq n^{1/3}$,

$$
\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T^{m'}_{4,1,2,4}-V_{4,1,2,4}(m')\right)_+\right]\leq Cn\times n^{10/3}\times n^{-\kappa_{b,4}^{1/4}2^{-3/4}}.
$$

Le choix de $\kappa_4=16^4/81\times 8\times(1/2)\|h\|^2\|\varphi_2^{(3)}\|_\infty^2\|\varphi_2'\|_\infty^2$ entraîne $\kappa_{b,4}\geq 16^4/81\times 8$, et donc $13/3-\kappa_{b,4}\times 2^{-3/4}\leq -1$. Ce qui termine la majoration de ce terme.

Pour le dernier terme ($l=1$), on doit effectuer une majoration un peu plus fine : on a,

$$
\begin{aligned}
\mathbb{E}\left[\left(T^{m'}_{4,1,2,1}-V_{4,1,2,1}(m')\right)_+\right] &= \left\|\varphi_2'\right\|_\infty^2 D_{m'}^3\\
&\quad\times\mathbb{E}\left[\left(6T^{m'}_{3,1,1}\left\|\hat{U}_n-id\right\|_\infty^2-\frac{\kappa_0\kappa_1}{\|\varphi_2'\|_\infty}\frac{D_{m'}\ln(n)}{n^2}\right)_+\right],\\
&\leq T^{m',a}_{4,1,2,1}+T^{m',b}_{4,1,2,1},
\end{aligned}
$$

avec

$$
\begin{aligned}
T^{m',a}_{4,1,2,1} &= \left\|\varphi_2'\right\|_\infty^2 D_{m'}^3\mathbb{E}\left[6T^{m'}_{3,1,1}\left(\left\|\hat{U}_n-id\right\|_\infty^2-\kappa_0\frac{\ln(n)}{n}\right)_+\right],\\
T^{m',b}_{4,1,2,1} &= \left\|\varphi_2'\right\|_\infty^2 D_{m'}^3\kappa_0\frac{\ln(n)}{n}\mathbb{E}\left[\left(6T^{m'}_{3,1,1}-\frac{\kappa_1}{\|\varphi_2'\|_\infty}\frac{D_{m'}}{n}\right)_+\right].
\end{aligned}
$$

Pour majorer le premier de ces deux termes, on majore grossièrement $T^{m'}_{3,1,1}$:

$$
\begin{aligned}
T^{m'}_{3,1,1} &= \sum_{j=1}^{D_{m'}}\left\{\frac{1}{n}\sum_{i=1}^{n}h(U_i)\varphi_j(U_i)-\mathbb{E}\left[h(U_i)\varphi_j(U_i)\right]\right\}^2,\\
&\leq \sum_{j=1}^{D_{m'}}4\left\|\varphi_j\right\|_\infty^2\|h\|_\infty^2,\\
&\leq 4\phi_0^2D_{m'}\|h\|_\infty^2.
\end{aligned}
$$

On est donc ramené au même type de majoration que ci-dessus :

$$
\begin{aligned}
T^{m',a}_{4,1,2,1} &\leq 4\phi_0^2\|h\|_\infty^2\left\|\varphi_2'\right\|_\infty^2 D_{m'}^4\mathbb{E}\left[\left(\left\|\hat{U}_n-id\right\|_\infty^2-\kappa_0\frac{\ln(n)}{n}\right)_+\right],\\
&\leq CD_{m'}^4\times n^{-\kappa_0},
\end{aligned}
$$

en utilisant le Corollaire 15. Ainsi, si $D_{m'}\leq n^{1/3}$,

$$
\sum_{m'\in\mathcal{M}_n}T^{m',a}_{4,1,2,1}\leq Cn\times n^{4/3}\times n^{-\kappa_0}.
$$

Le choix de $\kappa_0 = 10/3$ entraîne $7/3 - \kappa_0 \leq -1$. Ce qui termine la majoration de ce terme. Enfin, pour le dernier terme, on utilise le Lemme 30 qui prouve

$$\mathbb{E}\left[\sum_{m'\in\mathcal{M}_n}\left(6T_{3,1,1}^{m'} - V_{3,1,1}(m')\right)_+\right] \leq \frac{C}{n},$$

avec

$$V_{3,1,1}(m') = 6 \times 2(1+2\delta)\mathbb{E}\left[f^2(X_1)\right]\phi_0^2\frac{D_{m'}}{n}.$$

Si $\kappa_1 = \|\varphi_2'\|_\infty^2 \times 12(1+2\delta)\mathbb{E}\left[f^2(X_1)\right]\phi_0^2$, on a donc la majoration cherchée :

$$\sum_{m'\in\mathcal{M}_n} T_{4,1,2,1}^{m',b} = CD_{m_{\max}}^3\frac{\ln(n)}{n^2} \leq C\frac{\ln(n)}{n}.$$

$\square$

9.7.9. *Preuve du Lemme 35.*
On commence comme toujours par

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_{4,2,3}(m') - V_{4,2,3}(m')\right)_+\right] \leq \sum_{m'\in\mathcal{M}_n}\mathbb{E}\left[\left(T_{4,2,3}(m') - V_{4,2,3}(m')\right)_+\right],$$

et on raisonne pour m' fixé. On a montré en non adaptatif que

$$T_{4,2,3}^m \leq \|\varphi_2'\|_\infty\|\varphi_2''\|_\infty\frac{L^2}{\pi^2}D_m^2\left\|\hat{U}_n - id\right\|_\infty^3.$$

Par conséquent,

$$\begin{aligned}
&\mathbb{E}\left[\left(T_{4,2,3}(m') - V_{4,2,3}(m')\right)_+\right] \\
&\leq \|\varphi_2'\|_\infty\|\varphi_2''\|_\infty\frac{L^2}{\pi^2}D_{m'}^2\mathbb{E}\left[\left(\left\|\hat{U}_n - id\right\|_\infty^3 - \frac{\kappa}{\|\varphi_2'\|_\infty\|\varphi_2''\|_\infty L^2/\pi^2}\frac{\ln^{3/2}(n)}{n^{3/2}}\right)_+\right], \\
&\leq CD_{m'}^2 n^{-\kappa_b^{2/3}2^{-1/3}},
\end{aligned}$$

en utilisant le Corollaire 15 avec $p = 3$, et $\kappa_b = \frac{\kappa}{\|\varphi_2'\|_\infty\|\varphi_2''\|_\infty L^2/\pi^2}$. Ainsi, si $D_{m'} \leq n^{1/3}$,

$$\mathbb{E}\left[\max_{m'\in\mathcal{M}_n}\left(T_{4,2,3}(m') - V_{4,2,3}(m')\right)_+\right] \leq Cn \times n^{2/3} \times n^{-\kappa_b^{2/3}2^{-1/3}}.$$

Le choix de $\kappa = (8/3)^{3/2}2^{1/2}\|\varphi_2'\|_{\infty,[0;1]}\|\varphi_2''\|_{\infty,[0;1]}L^2/\pi^2$ entraîne $\kappa_b \geq (8/3)^{3/2}2^{1/2}$, et donc $5/3 - \kappa_b^{2/3} \times 2^{-1/3} \leq -1$. Ce qui termine la preuve.

$\square$

Références

[1] Antoniadis, A. ; Grégoire, G. ; Vial, P. Random design wavelet curve smoothing. *Statist. Probab. Lett.* 35 (1997), no. 3, 225-232.

[2] Baraud, Y. Model selection for regression on a random design. *ESAIM Probab. Statist.* 6 (2002), 127-146.

[3] Barron, A. ; Birgé, L. ; Massart, P. Risk bounds for model selection via penalization. *Probab. Theory Related Fields* 113 (1999), no. 3, 301–413.

[4] Birgé, L. Model selection for Gaussian regression with random design. *Bernoulli* 10 (2004), no. 6, 1039-1051.

[5] Birgé, L. ; Massart, P. Minimum contrast estimators on sieves : exponential bounds and rates of convergence. *Bernoulli* 4 (1998), no. 3, 329-375.

[6] Brunel, E. ; Comte, F. Penalized contrast estimation of density and hazard rate with censored data. *Sankhya* 67 (2005), no. 3, 441-475.

[7] Brunel, E. ; Comte, F. ; Guilloux, A. Nonparametric density estimation in presence of bias and censoring. *TEST* 18 (2009), no. 1, 166-194.

[8] Cai, T.T. ; Brown, L.D. Wavelet shrinkage for nonequispaced samples. *Ann. Statist.* 26 (1998), no. 5, 1783-1799.

[9] Comte F., Dedecker J., Taupin M-L., (2008). Adaptive density deconvolution with dependent inputs. *Mathematical Methods of Statistics* 17 : 87-112.

[10] Comte, F. ; Rozenholc, Y. A new algorithm for fixed design regression and denoising. *Ann. Inst. Statist. Math.* 56 (2004), no. 3, 449-473.

[11] Donoho, D.L. ; Johnstone, I.M. ; Kerkyacharian, G. ; Picard, D. Wavelet shrinkage : asymptopia ? With discussion and a reply by the authors. *J. Roy. Statist. Soc. Ser. B* 57 (1995), no. 2, 301-369.

[12] DeVore, R.A. ; Lorentz, G. Constructive approximation. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 303. *Springer-Verlag, Berlin*, 1993.

[13] Dvoretzky, A. ; Kiefer, J. ; Wolfowitz, J. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* 27 (1956), 642-669.

[14] Efromovich, S. Nonparametric curve estimation. Methods, theory, and applications. Springer Series in Statistics. *Springer-Verlag, New York*, 1999. xiv+411 pp. ISBN : 0-387-98740-1

[15] Fan, J. ; Gijbels, I. Variable bandwidth and local linear regression smoothers. *Ann. Statist.* 20 (1992), no. 4, 2008-2036.

[16] Gaïffas, S. On pointwise adaptive curve estimation based on inhomogeneous data. *ESAIM Probab. Stat.* 11 (2007), 344-364.

[17] Goldenshluger, A. ; Lepski, O. Bandwidth selection in kernel density estimation : oracle inequalities and adaptive minimax optimality. *Ann. Statist.* (to appear)

[18] Golubev, G. K. ; Nussbaum, M. Adaptive spline estimates in a nonparametric regression model. (Russian) *Teor. Veroyatnost. i Primenen.* 37 (1992), no. 3, 554-561 ; *translation in Theory Probab. Appl.* 37 (1992), no. 3, 521-529.

[19] Hardle, W. ; Tsybakov, A. Local polynomial estimators of the volatility function in nonparametric autoregression. *J. Econometrics* 81 (1997), no. 1, 223–242.

[20] Kerkyacharian, G. ; Picard, D. Regression in random design and warped wavelets. *Bernoulli* 10 (2004), no. 6, 1053-1105.

[21] Klein, T. ; Rio, E. Concentration around the mean for maxima of empirical processes. *Ann. Probab.* 33 (2005), no. 3, 1060-1077.

[22] Kohler, M. ; Krzyzak, A. Nonparametric regression estimation using penalized least squares. *IEEE Trans. Inform. Theory* 47 (2001), no. 7, 3054-3058.

[23] Lacour, C. Adaptive estimation of the transition density of a particular hidden Markov chain. *J. Multivariate Anal.* 99 (2008), no. 5, 787-814.

[24] Nadaraya E. On estimating regression. *Theory of Probability and its Application.* 9 (1964), 141-142.

[25] Pham Ngoc T-M. Regression in random design and Bayesian warped wavelets estimators. *Electron. J. Stat.* 3 (2009), 1084-1112.

[26] Tsybakov, A.B. Introduction à l'estimation non-paramétrique. Mathématiques & Applications (Berlin), 41. *Springer-Verlag, Berlin*, 2004.

[27] Watson G.S. Smooth regression analysis. *Sankhya Ser. A* 26 (1964), 359-372.